MICHIGAN STATE UNIVERSITY

Industrial/Organizational Psychology
and
Organizational Behavior

Team Decision Making in Hierarchical Teams:

A Final Report

John R. Hollenbeck and Daniel R. Ilgen

Michigan State University

MICHIGAN STATE UNIVERSITY
FOUNDED 1855

Michigan State University
East Lansing, Michigan 48824

Team Decision Making in Hierarchical Teams:

A Final Report

John R. Hollenbeck and Daniel R. Ilgen

Michigan State University

Prepared for:

Office of Naval Research

Grant Numbers:

N00014-93-1-0983

Modification A00001

Period Funded:

1 August 1993 - 31 October 1996

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
December, 1996

3. REPORT TYPE AND DATES COVERED

4. TITLE AND SUBTITLE
Team Decision Making in Hierarchical Teams: A Final Report

5. FUNDING NUMBERS
N00014-93-1-0983
Modification A00001

6. AUTHOR(S)
John R. Hollenbeck & Daniel R. Ilgen

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES)
Department of Management & Department of Psychology
Michigan State University
East Lansing, MI 48824

8. PERFORMING ORGANIZATION REPORT NUMBER
96-1

9. SPONSORING / MONITORING AGENCY NAMES(S) AND ADDRESS(ES)
Cognitive and Neural Sciences Division
Office of Naval Research
Arlington, VA 22217-5000

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12. DISTRIBUTION CODE
N00179

13. ABSTRACT *(Maximum 200 words)*

In the early 1990s, the authors began a series of research on decision making in hierarchical teams with distributed expertise. A simulation was developed and a model of decision making in these kinds of teams was proposed. The current research effort continued that work in two ways. First, work was done on the decision making model itself to replicate and revise it as needed in light of empirical findings. Second, research was conducted to investigate the effects of both static and dynamic factors on the ability of hierarchical teams to make accurate decisions. Static factors manipulated were the medium through which the teams interacted (computer mediated or face-to-face) and the architectures of the structures of information used to make decisions. Dynamically, one study investigated the decisions of teams operating in situations requiring sustained attention over a number of decision episodes. Another provided team decision making process feedback specifically designed to aid learning of key constructs in the model. The research on sustained attention and communications mode has been reported elsewhere. This final report will focus on support for the model itself, and the studies of architecture and team learning as influenced by process feedback. An appendix lists products produced during the funding period with citations to articles and presentations for the interested reader.

14. SUBJECT TERMS
Team decision making, distributed expertise, group dynamics, hierarchical teams

15. NUMBER OF PAGES
90

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT
Unclassified

18. SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

19. SECURITY CLASSIFICATION OF ABSTRACT
Unclassified

20. LIMITATION OF ABSTRACT

Team Decision Making in Hierarchical Teams:

A Final Report

For a variety of reasons, the use of teams in business, military, and medical contexts is on the rise. Although there is a great deal of research on group decision making, much of this research deals with people who are exposed to a common information base and who have to reach consensus on a single decision where there is no objectively verifiable answer (Ilgen, Major, Hollenbeck, & Sego, 1993). However, not all or even most important decision making groups are structured in this fashion. In many important decision making groups, all members are not equal in power and do not possess the same information. The decisions reached by the group also are not always consensus decisions. Instead, in organizations, decisions are often made in hierarchical decision-making teams with distributed expertise where the group or team leader has the final say on what the team decides.

In these types of teams, there are both status differences (leader vs. staff) and content area differences (staff members with different specialties), and the decision making team has to make a series of decisions where objectively verifiable feedback (in terms of being right or wrong) is available. Examples of these type of teams include managerial staffs, hospital emergency room teams, command and control teams, and academic research teams. Relative to the literature on jury decision making, there is far less in the way of programmatic research on hierarchical teams with distributed expertise, despite the fact that leader-staff (or judge-advisory) arrangements are ubiquitous in organizations (Brehmer & Hagafors, 1986; Sniezek & Buckley, 1995).

Hollenbeck, Ilgen, Sego, Hedlund, Major, and Phillips (1995) developed a theory of team decision making in hierarchical teams with distributed

expertise and reported two empirical studies supporting that theory. The research program of Hollenbeck and Ilgen that followed the initial development of the theory and a paradigm for studying team decision making was supported under the current grant. The purpose of the latter research was to investigate whether initial findings regarding the theory could be replicated and to extend the investigation of team decision making by investigating both static and dynamic factors affecting the decision making process. The general framework for static effects on team decision making involved manipulating critical conditions team settings, such as the complexity of decision objects, the medium through which intra team communication takes place, the time pressure under which decisions must be made, or the structure or architecture of information available to members. In all cases, such conditions were manipulated, and their effects on decision quality were assessed. The theory mentioned earlier was used as a guide for understanding the effects of these factors on the team decision making process.

In addition to the traditional static approach to decision making, later research also looked at decision making over time. In this case, mechanisms were introduced that aided teams in their ability to respond effectively to key constructs identified by the theory. Through these mechanisms teams were able to learn how to make better decisions.

The research that was conducted to extend our understanding of team decision making beyond the development of the initial model or theory has been reported in a number of different outlets. A complete listing of the work appears in Appendix A. This report will focus primarily on the last studies to be completed during this time period--ones that have not been written up for publication prior to this time. The two studies reported first replicate early work by (a) examining the degree to which the core constructs identified by

this theory explain performance differences between teams, and (b) examining the degree to which the influence of traditional variables studied in the group decision making literature are mediated by the core constructs identified by this theory. The research also extends previous research by (c) examining the impact of team structural variables on decisions, and (d) providing the first direct attempt to manipulate the core constructs of the theory to allow teams to learn team decision making processes predicted to improve the accuracy of their decisions. Before turning to these studies, we will briefly describe the theory on which they are based.

Overview of the Multilevel Theory

A detailed description of the Multilevel Theory of team decision making is provided in Hollenbeck et al. (1995, pp. 293-300). Briefly, this theory holds that team decision making accuracy is determined by constructs that occur at one of four levels: team, dyad, individual, and decision. The theory identifies the most critical variable at each of the three lower levels, and then forms aggregates of these variables at the team level in an effort to explain performance differences within and between teams in terms of the accuracy of team decisions.

The Multilevel Theory: Core constructs and propositions. At the decision level, the most critical variable is decision informity which is defined as the amount of relevant information the team, as a whole, has regarding any single decision. A team might have been well informed on one decision, but poorly informed on another. Characteristics of the decision context (e.g., time pressure) may affect the team's overall performance through their effect on this variable. Decision informity can be aggregated to form a team level construct referred to as team informity, which captures how well a team was informed, on average, across all the decisions it made. That is, even if one

holds the decision context constant, across a large number of decisions, some teams are likely to end up being better informed than others. The construct of team informity captures this aspect of the team decision process.

At the individual level, the most critical variable, according to this theory, is individual validity. It is defined as the degree to which any one staff member generates recommendations to the leader that are predictive of the "true" or "correct" decision for the team. Because the staff members are not making the final decisions, staff members need not make accurate decisions to be effective team members. That is, staff members may be biased (e.g., off by a constant). As long as the variance in their recommendations is related to variance in the "true" decision, the team leader in hierarchically structured teams can make effective use of their inputs by adjusting for the bias. Thus, staff members who might themselves make poor decisions if they were the team leader (because of their bias) may still be effective staff members. In teams with multiple staff members, individual team members' validities can be aggregated to form a team level construct referred to as staff validity. Operationally, staff validity has been simply defined as the average level of individual validity.

At the dyadic level, the most important variable identified by the Multilevel Theory is dyadic sensitivity. This variable reflects the degree to which the team leader effectively weighs each staff member's recommendation to arrive at the team's decision. For example, for any set of recommendations generated by a group of staff members, there is an ideal or optimal set of weights that can be applied to these recommendations to predict the correct decision. One can use policy capturing techniques to discern the weights that the leader places on these same recommendations when arriving at his or her final decision. The difference between the weight that the leader actually

assigns to a staff member's recommendation and the ideal or correct weight for that staff member's recommendation is used as the index of dyadic sensitivity (where a large difference implies low sensitivity). Again, in teams with multiple staff members, this variable can be aggregated to form a team level variable called hierarchical sensitivity, which captures the overall optimality of the leader's use of his or her staff.

These three concepts, decision informity, individual validity, and dyadic sensitivity, along with their team level analogs, constitute the core constructs of the Multilevel Theory. All other variables are referred to as non-core constructs. The major proposition of the Multilevel Theory, shown schematically in Figure 1, is that the core constructs are the most proximal causal factors that explain decision making accuracy, and that the effects of any noncore variables (e.g., variables traditionally considered important to group functioning) can be attributable to their influence on the core variables.

The Multilevel Theory: Empirical findings. Hollenbeck et al. (1995) report two studies that together employed approximately 125 four-person teams who spent roughly 1,500 hours making over 5,000 decisions on a simulated naval command and control simulator called TIDE^2. Results from the studies supported the theory and suggested that between 25% and 50% of the predictable variance in team decision making accuracy could be explained by the three core constructs. Noncore variables such as group cohesiveness, familiarity, experience, attrition, job knowledge, and role redundancy explained a smaller percentage of the variance. Most importantly, almost all of the effects for these latter variables was mediated by the three core constructs.

Similar results were obtained in Hedlund, Ilgen and Hollenbeck (1996) who examined the effect of communication media (face-to-face versus computer

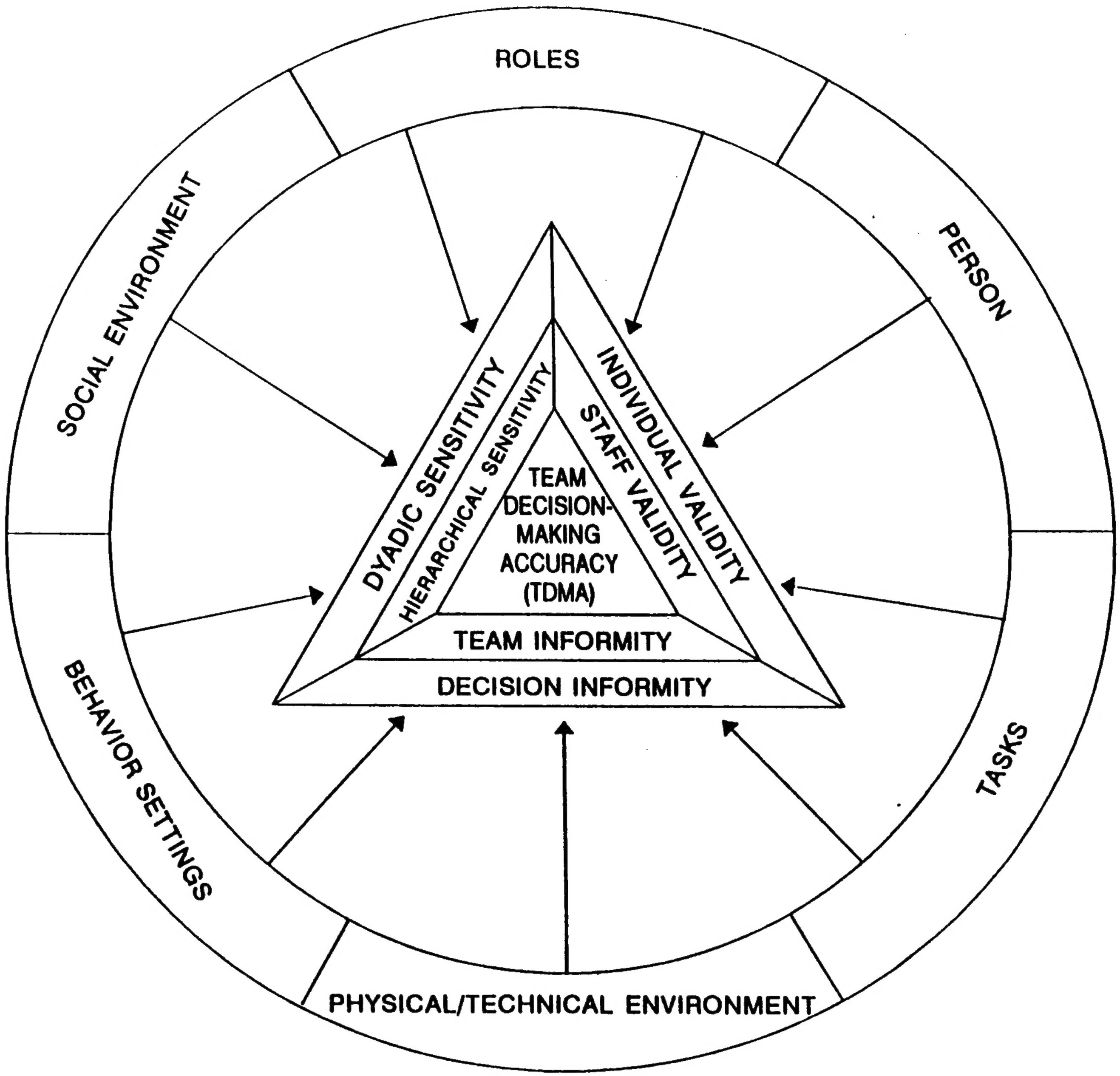


Figure 1. Overview of the multilevel theory of team decision making

mediated) on decision making accuracy in 64 four-person teams. In this study, the core constructs identified by the Multilevel Theory explained 43% of the variance in team decision accuracy, and all the effects for communication media (which accounted for 8% of the variance when considered alone) were mediated by the core constructs.

In all of these three studies reported elsewhere, the value of decomposing overall team decision making accuracy into the core constructs could be seen when the impact of some non-core construct had multiple, but counter-acting, influences on different core constructs. For example, although face-to-face teams slightly outperformed computer mediated teams in Hedlund et al. (1996), the overall impact of communications medium if treated only as a main effect was under-estimated, because it had counter-acting influences on the core constructs. Part of the reason face-to-face teams outperformed computer mediated teams in that study was that face-to-face teams were better informed and developed higher staff validity (which enhances team performance). On the other hand, the face-to-face teams, relative to computer mediated teams, performed worse in terms of hierarchical sensitivity. Leaders who simply received cold, detached recommendations were better able to learn how to weigh these recommendations than were leaders who actually observed the staff discussions. This suggests that an ideal arrangement might be one where the staff meets face-to-face, but then sends the recommendations on to a leader that did not participate or observe the discussion. Similar types of complex, counteracting influences have also been found for variables like team member familiarity and role redundancy. That is, relative to the core constructs, these variables have both good and bad aspects to them, but these effects are masked when one simply examines overall decision accuracy (Hollenbeck et al., 1995).

The research reported here replicates and extends previous research on team decision making in hierarchical teams. It is a replication in the sense that it assesses the predictive value and mediating role of the same three core constructs of the Multilevel Theory introduced in Hollenbeck et al. (1995). Study 1 is an extension in that it explores the mediating role of the core constructs with respect to two additional "noncore" variables--group architecture and implicit coordination. These two variables are increasingly cited as being critical factors in the group decision making literature, and thus linking the core constructs to these variables illustrates how this theory relates and contributes to the existing knowledge base.

Study 2 is an extension in that it is the first study where team members are provided with direct feedback in terms of the core constructs. That is, previous studies have manipulated the noncore constructs and observed naturally occurring variation on the core constructs. In Study 2, we randomly assigned teams to conditions where feedback on the core constructs was provided to half the teams but not the others.

Study 1

Implicit Coordination and Group Architecture

Implicit coordination. There exists considerable support for the position that team members need to construct a shared mental model regarding each member's capacities and needs (Klimoski & Mohammed, 1994). Teams that succeed in developing effective shared mental models have a common understanding of the team's task and each member's role in accomplishing the task (Cannon-Bowers, Salas, & Converse, 1990; Rouse, Cannon-Bowers, & Salas, 1992). One of the primary benefits of common team mental models is that the shared understanding provided by such frameworks obviates the need for a great deal of communication aimed toward coordinating actions. Freed of the need to

insure that others know what to do and when to do it, team members can concentrate more on their own individual responsibilities (Orasanu, 1990). This latter aspect of shared mental models has been referred to as implicit coordination. The hallmark of teams that have good implicit coordination is that members' needs are met without a great deal of direct, explicit communication.

From a Multilevel Theory perspective, teams with good implicit coordination should be able to transmit critical information to each other faster and more efficiently (thus achieving high levels of team informity). Because information is received in a timely manner and team members are focused on task performance rather than requesting missing information, we would also expect this to enhance the validity of staff members' recommendations, as well as the appropriateness of the leader's weighting strategy.

Group architecture. As we noted earlier, the current program of research is interested in teams structured to have a leader with a staff. Leader-staff arrangements can take on many different forms. This study seeks to explore different types of team architectures and the role played by implicit coordination in these different architectures. We define team architecture as the means by which information and expertise is distributed throughout the team (Carley, 1991, 1992; Carley & Lin, 1995). Architecture is a complex, multifaceted construct that deals with group structure, information distribution, and communication requirements. In an attempt to promote better decision making team architectures or structures are often manipulated (Rogelberg, Barnes-Farrell, & Lowe, 1992).

The number of different architectures that could be configured for a set number of individuals (e.g., four people) attempting to cover some fixed area

of decision space (e.g., a decision that needs to be made based on a set of nine pieces of information) is vast. Our primary interest is not to map a large number of possible architectures then compare and contrast them (a task that may be more suitable to mathematical modeling approaches--see, for example, Carley & Lin, 1995). Rather, our interest is in sampling a small number of different architectures that differ in interesting ways. We will then compare the degree to which the core constructs of the Multilevel Theory mediate the impact of the structures on team decision making accuracy.

Three different team architectures were examined. One set of teams was comprised of staff members who were generalists. In the generalist teams, each staff member can access all of the available information but has only general knowledge in terms of interpreting that information. Since each person has access to all the available information, there is minimal need for communication or coordination among the team members. For contrast, we examined staffs composed of specialists. Specialist teams have staff members who possess highly precise knowledge in terms of interpreting information, but each staff member has access to only a few pieces of information. Since each staff member can only cover a small portion of the entire decision space, these teams required some degree of coordination to ensure each staff member had an accurate cognitive representation of the decision object as a whole. Finally, a third set of teams were configured as general-specialists. Their members had range and precision at levels intermediate to those associated with the other two conditions. Figures 2a through 2c schematically depict the three types of structures examined in this study.

Costs and benefits are associated with each team structure. The broad range of the generalists' knowledge means that each staff member can operate

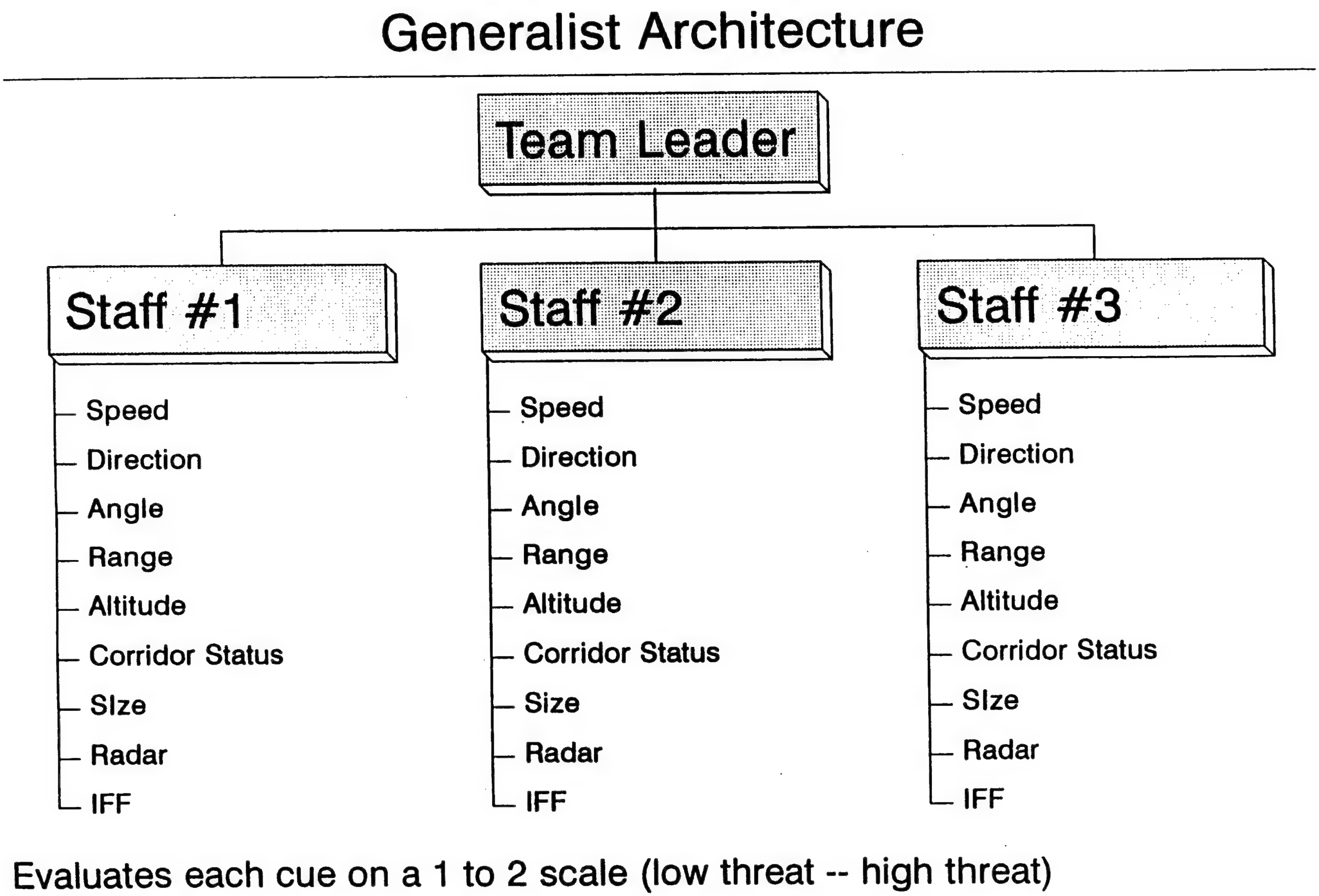
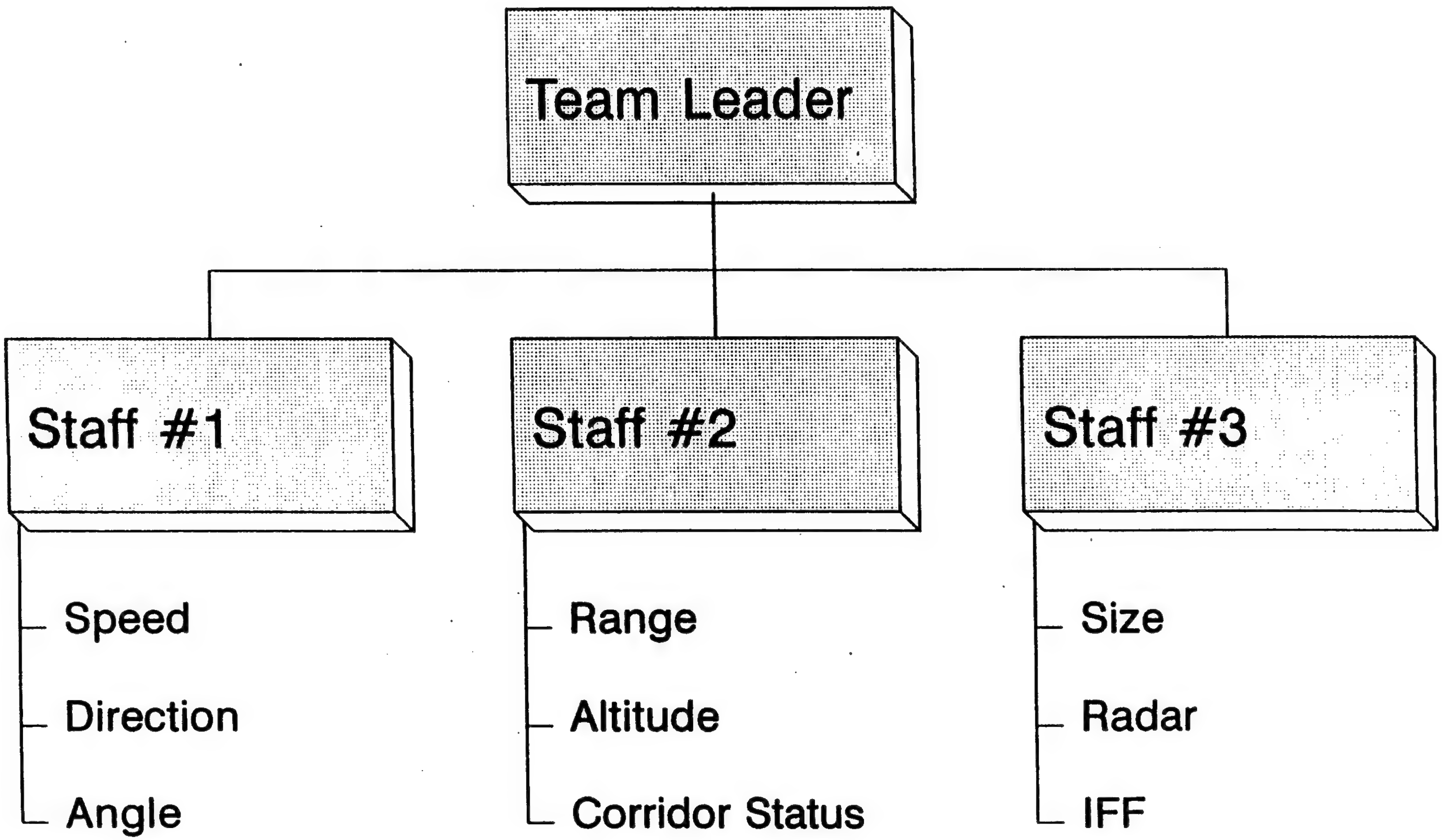


Figure 2a. The generalist architecture

Specialist Architecture



Evaluates each cue on a 1 to 5 scale (low threat -- high threat)

Figure 2b. The specialist architecture

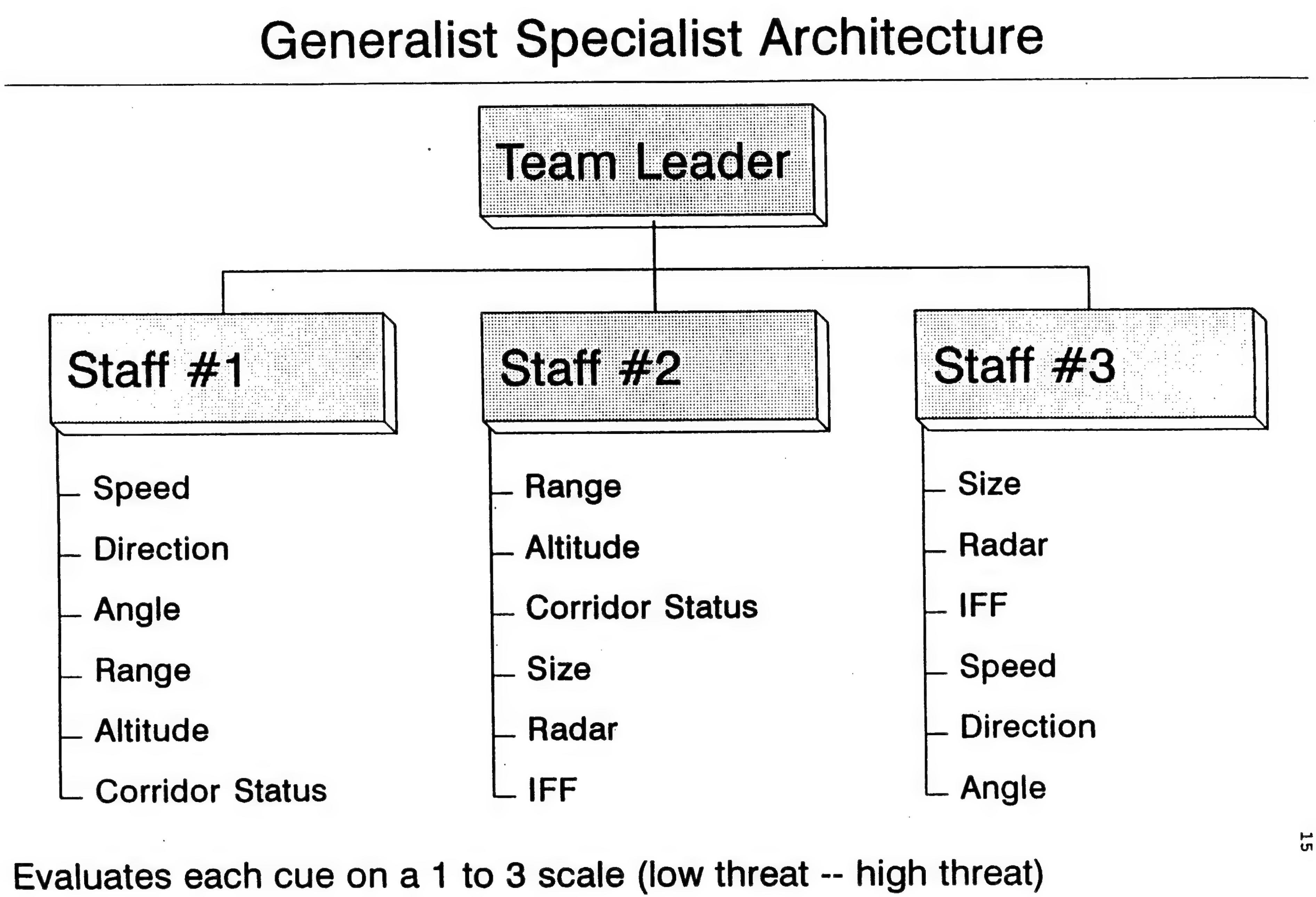


Figure 2c. The general-specialist architecture

independently from the other staff members. On the other hand, the generalists can make only gross distinctions on each dimension. The specialists and general-specialists can make finer distinctions within dimensions relative to the generalists, but face demands for coordination because they cannot cover the entire decision space.

Specialists, generalists, and general-specialists represented three architectures that were manipulated independently. However, their effectiveness with respect to performance was expected to interact with the quality of the team's coordination behavior. Therefore, coordination was measured and teams were classified as either good or poor on it (see "Research Design" for further detail).

Figure 3 shows a summary of the predicted levels of performance for the five types of teams examined here. In general, our prediction is that specialist and general-specialist teams that are well coordinated (i.e., WCSP and WCGS in Figure 3) will outperform generalists teams (i.e., Gen in Figure 3) due to their increased precision. However, poorly coordinated specialist and general-specialist teams (i.e., PCSP and PCGS in Figure 3) will perform worse than generalists teams because of their inability to gain timely access to information.

In addition, in line with the previous literature on the Multilevel Theory, we predicted that any differences in team performance attributable to architecture or implicit coordination will be explained by the three core constructs specified by the theory. For example, the generalists, because of their access to all information, should generate higher levels of decision informity relative to the other types of staffs. However, the gross level of knowledge on the part of the generalists may limit their individual validity relative to the more finely grained specialists. Finally, because they all

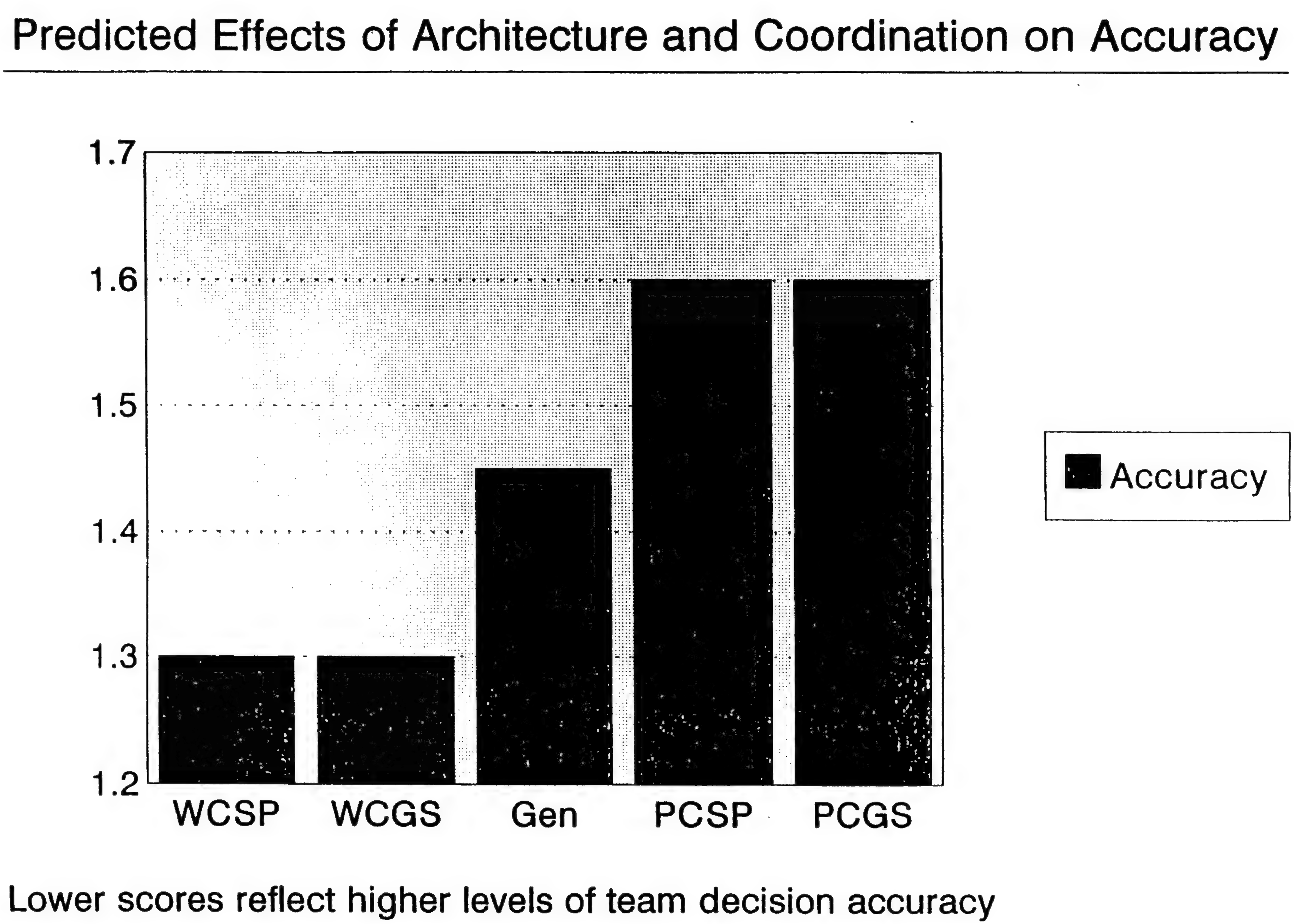


Figure 3. Predicted effects of architecture and implicit coordination on team decision accuracy

focus on the exact same information, the recommendations of the generalists will tend to be redundant. The opinions of the specialists on the other hand, are likely to be more varied. This variability increases the need for developing differentiated weighting systems with specialized staff members, in order to reap the benefits of "constructive controversy" (Tjosvold, 1982). Thus, hierarchical sensitivity is likely to be considered more critical (and therefore receive more attention) to leaders of specialized staffs relative to generalists.

Method

Participants. Research participants were 420 undergraduate students at a large midwestern university who were arrayed into 105 four-person teams. In return for their participation, each received course credit and was also eligible to earn a cash prize based upon team performance.

Task. Since this was meant to be a replication of Hollenbeck et al. (1995), the same task (i.e., $TIDE^2$) and procedures were used in this study as that one (see pages 301-303 of Hollenbeck et al., 1995 for more details). $TIDE^2$ is a software program for a decision task simulation that presents participants with values on a number of attributes of a problem or object.

In this particular study, $TIDE^2$ was programmed to simulate a naval command and control scenario with a leader and three staff members. Each team member sat at a work station in front of a computer that was networked to all other team members. The team's task was to monitor the airspace surrounding the team. When any aircraft came into this airspace, each team member needed to gather some information about particular attributes of the aircraft (e.g., its speed, direction, angle, range, size, etc.), and then arrive at a judgment regarding the appropriate response to make toward the aircraft. Judgments and decisions were rendered on a seven-point continuum that varied in

aggressiveness from Ignore (the lowest level of aggressiveness), to Defend (the most aggressive response). Intermediate anchors on this scale were arranged in increasing levels of aggressiveness.

The team's decision was made by the leader, after receiving input from the staff. Once made, the leader's response was compared to the correct decision. This correct decision was based on translating the rules of the simulation exercise into a linear combination of the attributes and applying the equation to the attribute values of the stimulus aircraft.

Measurement of team decision making accuracy. Accuracy was defined as the degree to which the leader's decision matched the correct decision. The performance scale and the verbal anchors associated with it were: Hit for no difference, Near Miss for a decision one off from the correct decision, a Miss for being two off, an Incident for being three off, and a Disaster for a difference of four or more. Following the leader's decision, each person in the team received performance feedback that told the team's performance on the trial. There were 36 trials (i.e., aircraft to evaluate).

Research design. This study is best conceptualized as a nonfully crossed 3 X 2 Design, where there were three different team architectures (generalist, general-specialist, specialist) and two levels of team coordination (well coordinated, poorly coordinated). The design is not fully crossed because there is no demand for coordination on the part of the generalists (who by definition, can obtain all the needed information on their own). Thus, this leaves five groups that can be compared, generalists (Gen), well coordinated specialists (WCSP), poorly coordinated specialists (PCSP), well coordinated general-specialists (WCGS) and poorly coordinated general-specialists (PCGS). In this study, the generalists serve as a control group for the other combinations of coordination and architecture.

Measurement of implicit coordination. As we noted earlier, the hallmark of teams with good implicit coordination is that each team member obtains what is needed from other team members without having to directly ask for help. The TIDE^2 program records all responses of each team member, and also aggregates these into indices that capture the level of implicit team coordination. Figure 4 shows the steps necessary to successfully execute a question and answer with the TIDE^2 simulation task. That is, if Team Member A needs some information from Team Member B, he or she needs to send a query from A to B, this query then needs to be received by B, who must then transmit the information to A, who must then receive the information, thus completing a four-step sequence.

Implicit coordination is defined here in terms of "wasted motions," that is, behaviors that take time and effort but fail to result in any information exchange or gain. Teams that generated a large number of wasted motions in moving information around the system were viewed as lacking implicit coordination.

Wasted motions could take on four different forms as shown in Figure 4. If Team Member A queries Member B, but B fails to receive this query, this results in one wasted motion (hereafter referred to as a "Slight"). If Team Member B receives this query (i.e., goes to his message center and reads it) but fails to do what is asked of him, this results in two wasted motions (hereafter referred to as being "Unresponsive"). If B receives the query and answers it, but the person who initiated the communication fails to go back and receive this information by looking at it when notified that someone has responded to his request, the result is now three wasted motions (hereafter referred to as being "Forgetful"). Finally, a fully completed four step sequence that leads to information exchange (hereafter referred to as a

Indicators of Poor Implicit Coordination

$$\boxed{\text{Query A to B}} \xrightarrow{\#1} \boxed{\text{Receive B from A}} \xrightarrow{\#2} \boxed{\text{Transmit B to A}} \xrightarrow{\#3} \boxed{\text{Receive A from B}}$$

Label	Nature of Failure	Number of Wasted Motions
Slight	Query not Received (no #1)	1
Unresponsive	Query Received but not responded to (#1 but no #2)	2
Forget	Query Received and Responded to but not Re-received (#1 and #2 but no #3)	3
Learn	Completed Sequence (#1, #2, and #3)	2

Figure 4. Indicators of poor implicit coordination

"Learn") generates two wasted motions. Two motions are wasted in the sense that a team with perfect implicit coordination can get the same result with two steps. That is, if team members know each others' information needs, there should be no need to send and receive queries--the information should just be sent and received in two steps (hereafter referred to as a "Lecture"). Because the teams could communicate with each other via open text messages and were given time prior to the experimental trials to develop a task strategy, a well coordinated team should have been able to avoid direct queries during real-time task engagement.

In this study, the total sum of wasted motions was a measure of implicit coordination (high score indicates poor implicit coordination). The four types of wasted motions formed a unidimensional scale with a coefficient alpha of .74 implying that they tapped a common underlying construct; teams that displayed one type of wasted motion tended to display others.

Since generalists teams did not need to communicate, the measure was not relevant to these teams. However, the two nongeneralists types of teams were divided into well coordinated and poorly coordinated groups by splitting them at the median, thus creating the five different types of teams listed in Figure 2.

<u>Manipulation of team architecture</u>. Teams' were randomly assigned to one of three architectures: generalist, specialist, or general/specialist. In the generalist architecture, staff members could measure all attributes of aircraft (nine in all), but they only had a general idea of the meaning of information from each dimension. That is, they were taught only one "critical value" on each of the nine attributes, and this critical value broke up the dimension into two gross categories (threatening or nonthreatening). For example, on the dimension of altitude, they were trained that any aircraft

under 21,000 feet was threatening, but aircraft above this altitude were nonthreatening.

Staff members in generalist teams also had to learn all three "rules of engagement." These rules specified how values on specific dimensions combined to determine the overall nature of the threat associated with a specific aircraft. For example, one of the rules of engagement (i.e., the location rule) specified a three-way interaction between altitude, range and corridor status, such that an aircraft had to be threatening on all three attributes (low, close and outside the corridor) before it was considered threatening overall.

In the specialist condition, staff members could measure only three attributes but were taught four critical values on each attribute. These four critical values divided each dimension into five precise categories (very threatening, somewhat threatening, uncertain, somewhat nonthreatening and very nonthreatening). The critical values were the actual cut-offs used to generate the "true" score for each aircraft, and thus had higher fidelity than the system used by the generalists. For example, on the dimension of altitude, specialists were trained that any aircraft under 13,000 feet was very threatening, 14,000 to 16,000 feet was somewhat threatening, 17,000 to 23,000 feet was uncertain, 24,000 to 26,000 feet was somewhat nonthreatening and 27,000 to 35,000 feet was not at all threatening. Finally, staff members in specialist teams only had to know one rule of engagement.

In the general-specialist condition was a mix of the previously described two in which each staff member could measure six attributes, but was taught only two critical values on each. These two critical values broke up each dimension into three categories (threatening, uncertain, nonthreatening).

Each staff member in general-specialist teams had to know two rules of engagement.

Note that the three group architectures are roughly equivalent in their information storing and processing demands. Each staff member in every condition had to store approximately the same number of points in working memory, but the information possessed by each differed in terms of breadth, depth, and specialization.

Both Specialists and Generalist-Specialists needed to exchange information with other team members because no one staff member ever possessed all the information necessary for evaluating the aircraft on the rule(s) for which he/she was responsible. That is, if the staff member was responsible for the location rule, which dealt with altitude-range-corridor status, he or she could only measure one of these attributes directly. The remaining two pieces of information had to come from other team members, hence creating interdependence and a need for coordination.

Measurement of the core constructs. The three core constructs were measured in the same manner as in Hollenbeck et al. (1995; see pp. 296-298 for more details). Briefly, team informity was simply the total amount of information available to the team when they made their decision averaged across all trials; staff validity was the average correlation between the staff members' recommendation and the correct decision; and hierarchical sensitivity was the average of the absolute differences between the weights actually assigned by the leader to his or her staff's recommendations and the ideal set of weights obtained by regressing the true decision on the staff members' recommendations. The level of analysis for these constructs was the team level, and hence the sample size was 105 for all of these measures.

We also measured analogs of these constructs at levels below that of the team. For decision informity, the lowest level of analysis was that of the trial. Each aircraft passing through the airspace represented a trial requiring a decision. From now on we shall refer to trial level decisions as the decision trial. The number of decisions made across all teams was 3,780 (i.e., 105 teams that each made 36 decisions). For individual validity, the level of analysis is at the individual staff member level, and, therefore, the number of observations was 315 (i.e., 105 teams that each contained three staff members). Finally, for dyadic sensitivity, the level of the analysis is the vertical leader-staff dyad for which there was also 315 decisions (i.e., 105 teams that each contained 3 leader-staff dyads).

Study 1 Results and Discussion

Table 1 shows the descriptive statistics associated with the variables measured in this study at the team level (n = 105).

The impact of the core constructs. The results of the first stage of data analysis are shown in Table 2 where team decision accuracy is regressed on the three core constructs. As a set, the three core constructs explained a statistically significant 42% of the variance in team decision making accuracy. There was also a statistically significant effect for each of the three core constructs, although in terms of effect size, the majority of this variance was explained by staff validity (incremental R^2 of .28) relative to team informity and hierarchical sensitivity (incremental R^2's of .04 and .10). Interactions among the core constructs were tested, but none were statistically significant, and these were omitted from Table 2 to conserve space.

Table 1

Descriptive Statistics for All Variables at the Team Level (n = 105)

	Mean	SD	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) Team Decision Accuracy	1.50	.26	---							
(2) Team Informity	.64	.14	-.20	---						
(3) Staff Validity	.32	.14	-.56	.33	---					
(4) Hierarchical Sensitivity	.20	.11	.47	.01	-.29	---				
(5) Well Coordinated Specialists	.17	.38	-.16	-.48	.23	-.13	---			
(6) Poorly Coordinated Specialists	.15	.36	.00	-.33	-.05	-.17	-.19	---		
(7) Well Coordinated GenSpecs	.15	.36	.17	-.19	-.27	.06	-.19	-.18	---	
(8) Poorly Coordinated GenSpecs	.17	.38	.32	.08	-.33	.22	-.21	-.19	-.19	---

$r > .16$; $p < .05$

Note. Decision accuracy and hierarchical sensitivity are deviation scores and hence low values reflect high standing on the constructs.

Table 2

The Effects of the Core Constructs on Team Decision Accuracy (n = 105)

Hierarchical Step	Independent Variable	Total R^2	Incremental R^2
(1)	Team Informity	.04*	.04*
(2)	Staff Validity	.32*	.28*
(3)	Hierarchical Sensitivity	.42*	.10*

*denotes $p < .05$

To illustrate the nature of these effects, Figure 5 shows the difference in decision accuracy and disaster rates (where a disaster is defined as a highly erroneous decision that is off by four or more points) for teams that were high on all the core constructs versus teams that were low on all the core constructs. Decision accuracy was higher for teams that were high on all three constructs by almost two standard deviation units, and the disaster rates for teams low on all three core constructs were twice as high as the disaster rates for teams high on all three.

The impact of the noncore constructs. Table 3 shows the results for the second stage of analysis where team decision accuracy is regressed on four dummy variables that capture the five types of teams (where the generalist teams are used as the control group). As a set, the nature of the team accounted for 17% of the variance in team decision accuracy. Figure 6 illustrates these results. On the one hand, as predicted, both poorly coordinated nongeneralists teams were outperformed by the generalist teams. However, the well coordinated specialist team performed at the same level as the generalists team, and the well coordinated general-specialists actually performed worse than the generalists.

Mediating role of the core constructs. The third stage of the analysis tests the degree to which the effects for the noncore variables (architecture and implicit coordination) are mediated by the core variables. Table 4 shows the results of regressing team decision accuracy on the dummy variables capturing team type, after controlling for the core constructs. The results show that after entering the core constructs, the variables capturing the nature of the team explain no variance in team decision accuracy. This implies that all the effects of team type on team performance are mediated by the core constructs.

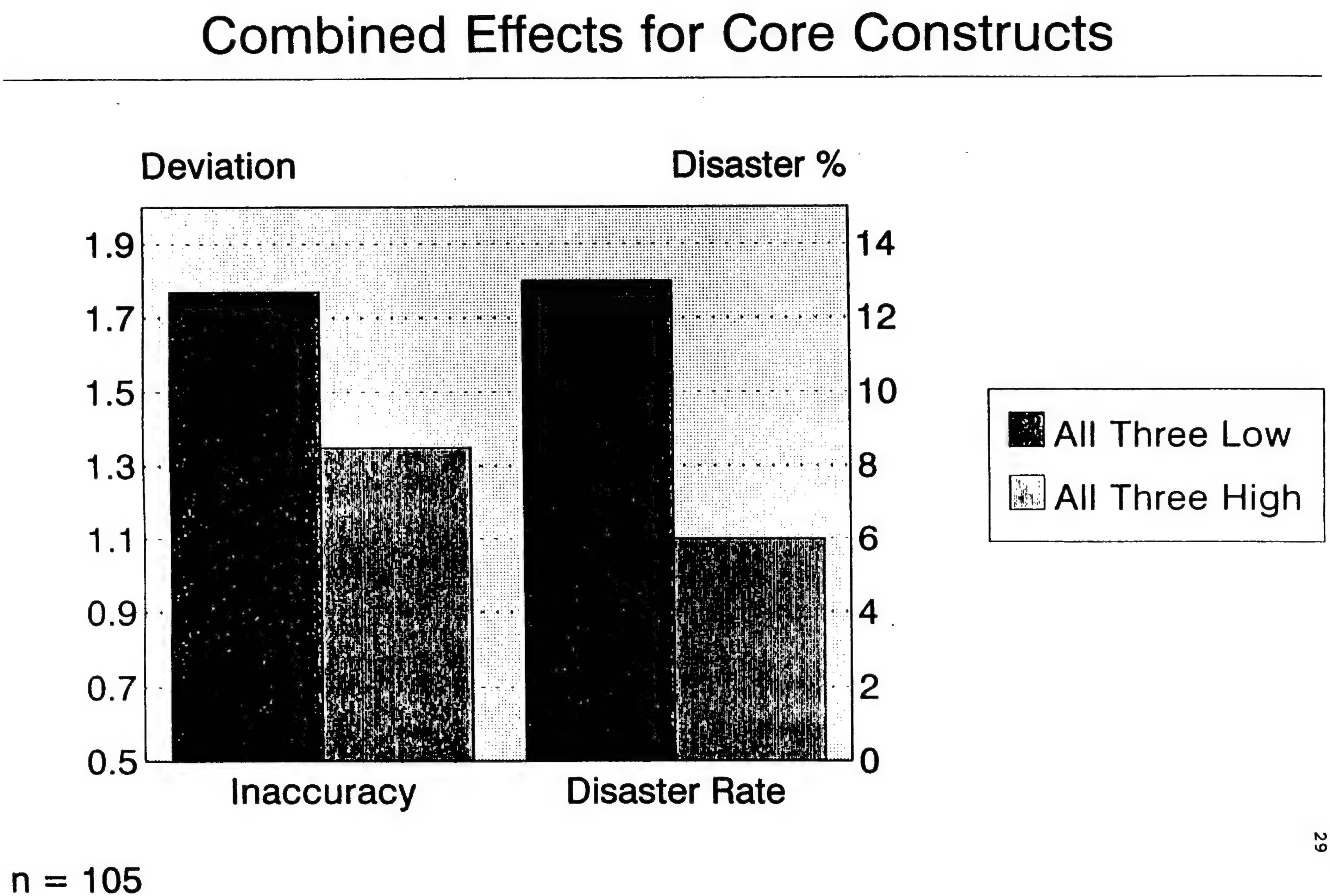


Figure 5. Combined effects for the three core constructs on team decision making accuracy and disaster rates

Table 3

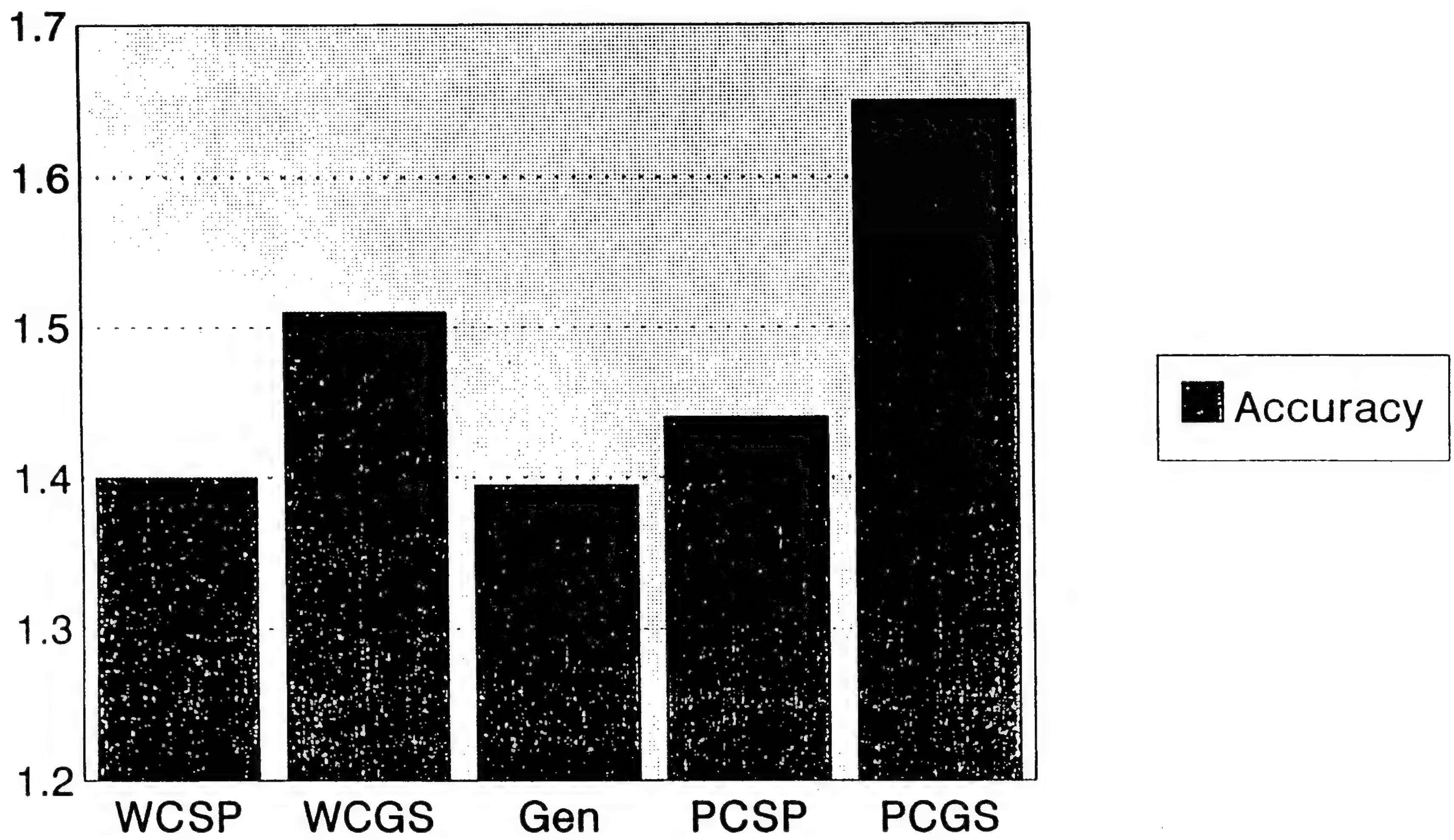
The Effects of Team Structure and Implicit Coordination on Team Decision Accuracy

Hierarchical Step	Independent Variable	Total R^2	B Weight
(1)	Well Coordinated Specialists		.00
	Poorly Coordinated Specialists		.10
	Well Coordinated GenSpecs		.20*
	Poorly Coordinated GenSpecs	.17*	.28*

* denotes $p < .05$

Note. B weights reflect mean differences between the control group (Generalists) and the group specified as the variable.

Effects of Architecture and Coordination on Decision Accuracy



Lower scores reflect higher levels of team decision accuracy

Figure 6. Effects of architecture and implicit coordination on team decision making accuracy

Table 4

The Mediating Role of the Core Constructs

Hierarchical Step	Independent Variable	Total R^2	Incremental R^2
(1)	Team Informity	.04*	.04*
(2)	Staff Validity	.32*	.28*
(3)	Hierarchical Sensitivity	.42*	.10*
(4)	Well Coordinated Specialists		
	Poorly Coordinated Specialists		
	Well Coordinated GenSpecs		
	Poorly Coordinated GenSpecs	.42*	.00

*denotes $p < .05$

Nature of the mediating effects. To test the nature of these mediating effects, we treated the core constructs as dependent variables and examined the degree to which each was affected by either team architecture or coordination. Since variability in the core constructs exists both within and across teams, we used repeated measures regression (Cohen & Cohen, 1983; Hollenbeck, Ilgen, & Sego, 1994) to decompose the nature of these effects and partition the variance. That is, the manipulations and measures of the independent variables are at the team-level, but some of the variance in the core constructs (i.e., the dependent variables) is at lower levels (i.e., the decision-level, individual-level or dyad-level).

Repeated measures regression allows one to partition the variance attributable to different levels, thus providing a more sensitive test of the effects of independent variables that can operate only at one level. So for example, the architecture of the teams was manipulated at the team level, and therefore any within team variance in staff member validity cannot possibly be explained by this manipulation. Thus, this within team variance needs to be isolated and removed prior to testing for the effects of the architecture manipulation.

Figures 7a, 7b and 7c show the results of this variance partitioning for each of the core constructs. For decision informity, 63% of the variance in this variable was due to between team variance (i.e., team informity), whereas 37% was attributable to variance across decisions within teams. The four variables capturing the teams architecture and level of coordination explained 37% of the total variance in informity, which translates into 59% (i.e., 37%/63%) of the between teams variance, which is statistically significant ($F = 35.98$; $df = 4, 100$; $p < .05$). The nature of these effects is shown in

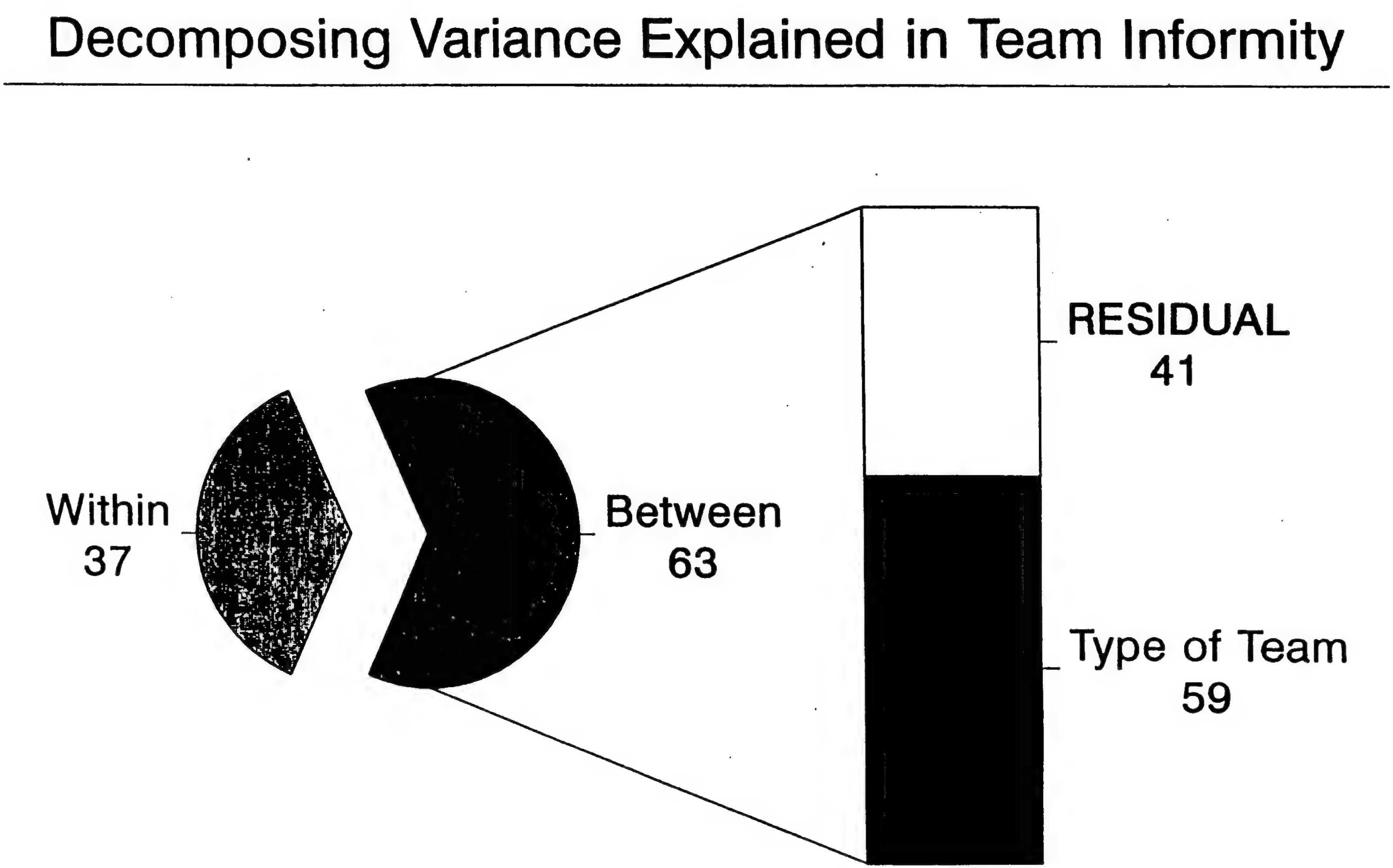


Figure 7a. Decomposing variance in team informity

Decomposing Variance Explained in Staff Validity

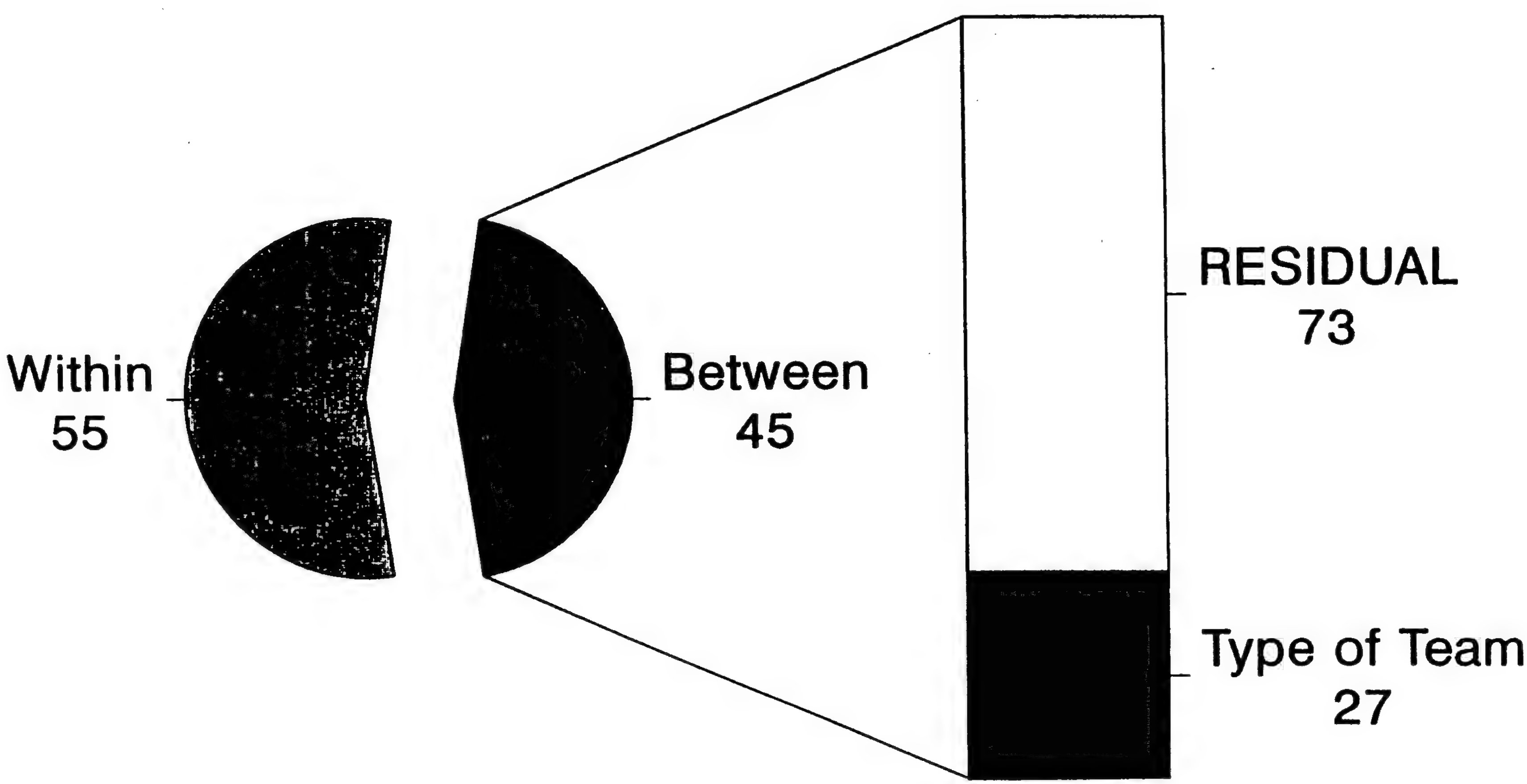


Figure 7b. Decomposing variance in staff validity

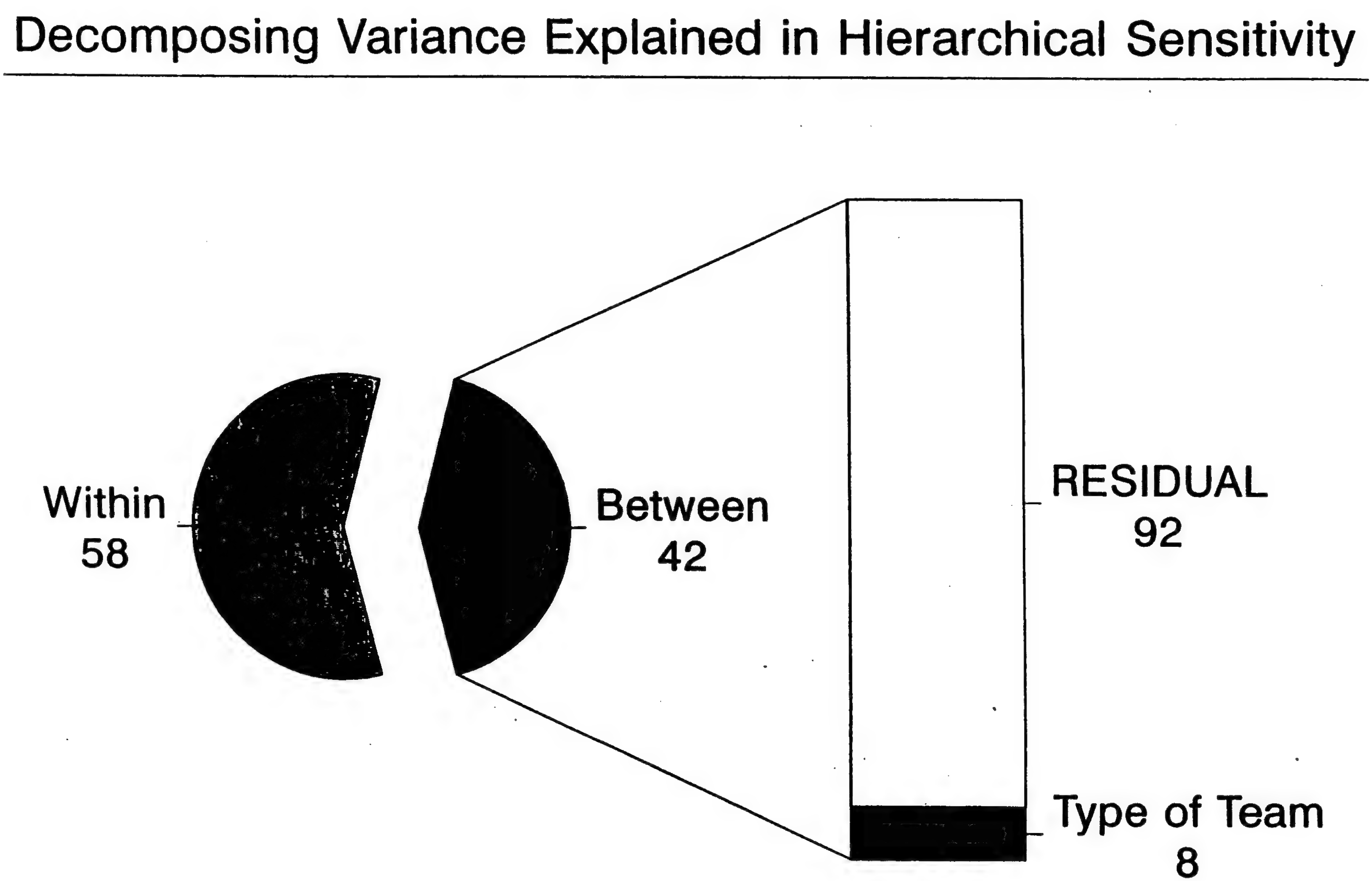


Figure 7c. Decomposing variance in hierarchical sensitivity

Figure 8a. Generalist teams were much more informed than any of the specialist or general-specialist teams.

Although we will return to this point later in the general discussion, what should be emphasized at this point is that regardless of coordination, the specialist teams had the least amount of information to work with relative to the generalists. Yet, looking back at Figure 6, these teams performed just as well as the generalists when they were well coordinated. Thus, the information requirements for individual staff members in teams with this type of architecture are vastly reduced without any loss in decision accuracy.

For individual validity, 45% of the total variance in this variable was due to between team variance (i.e., staff validity), whereas 55% was attributable to variance across staff members within teams. The four variables capturing the teams' architecture and level of coordination explained 12% of the total variance in staff validity, which translates into 27% (i.e., 12%/45%) of the between teams variance, which is statistically significant ($F = 12.33$, $df = 4, 100$; $p < .05$). The nature of these effects is depicted in Figure 8b, where it can be seen that the recommendations of staff members in generalist teams were higher in validity compared to all the other groups except the well coordinated specialists. This latter group achieved comparable levels of validity despite possessing a great deal less information.

For dyadic sensitivity, 42% of the total variance was due to between team variance (i.e., hierarchical sensitivity), whereas 58% was attributable to variance across the three leader-staff dyads within teams. The four variables that captured team architecture and level of coordination explained 3.5% of the total variance in hierarchical sensitivity, which translates into 8% (i.e., 3.5%/42%) of the between teams variance, which is statistically

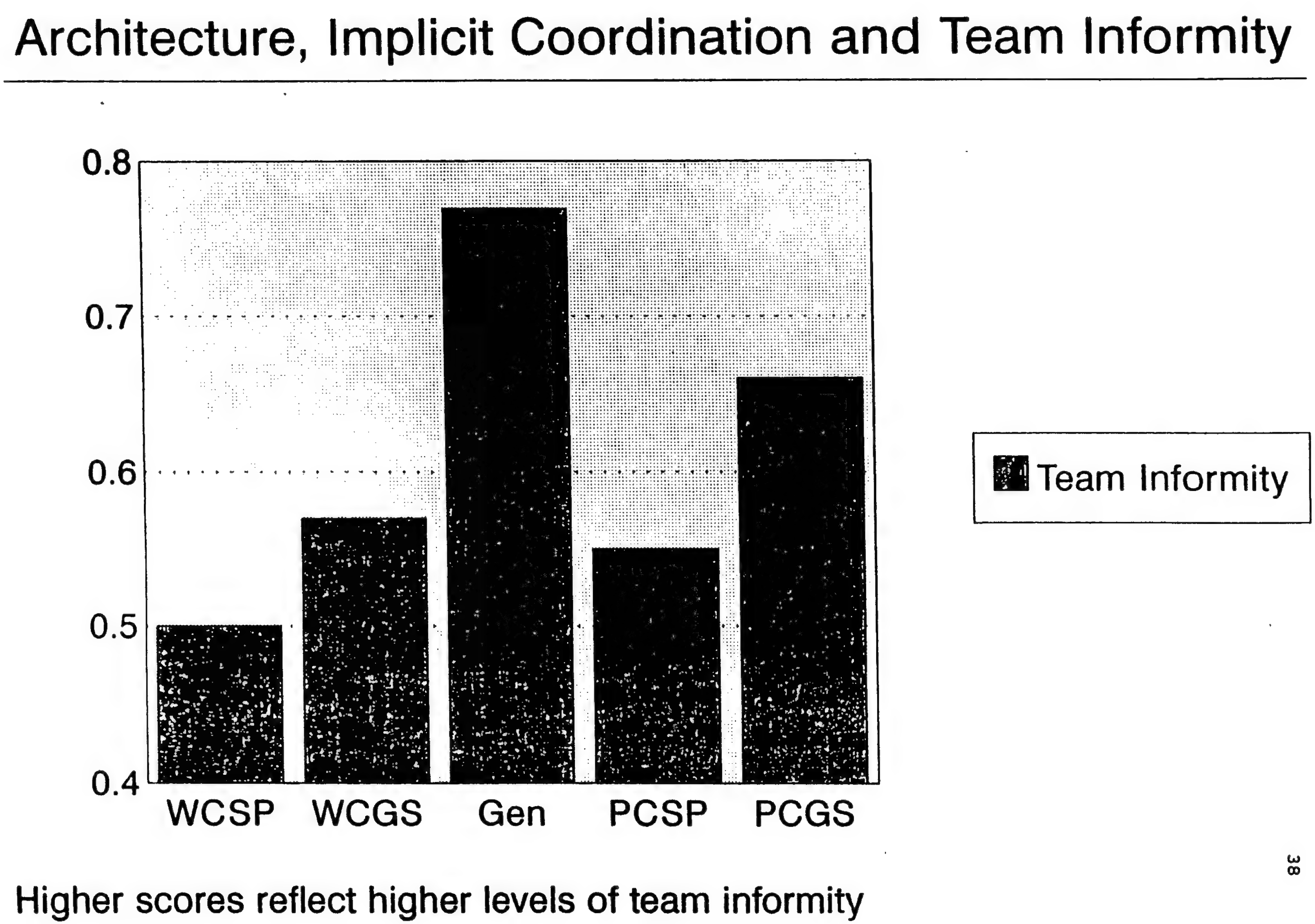
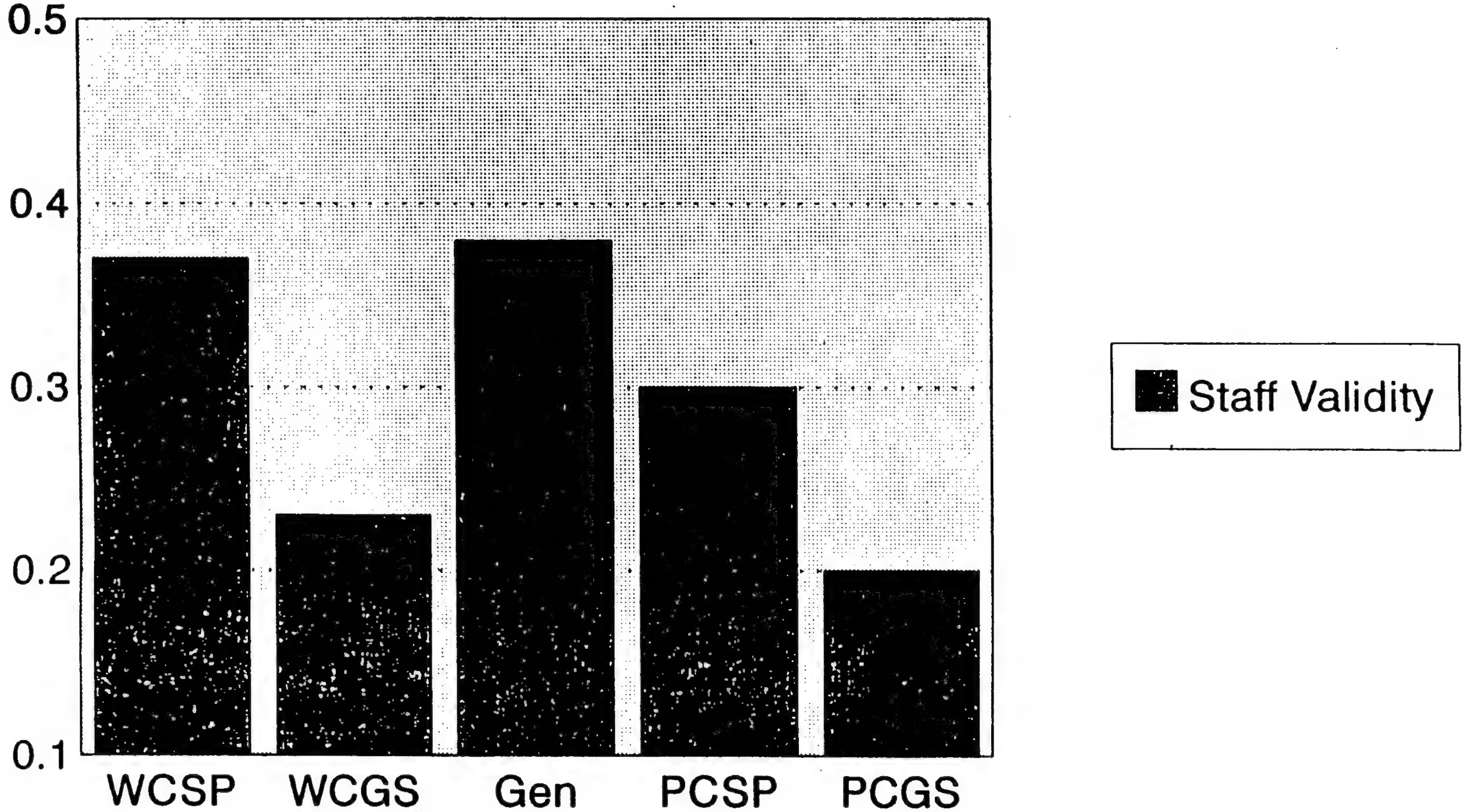


Figure 8a.. Architecture and implicit coordination on team informity

Architecture, Implicit Coordination and Staff Validity



Higher scores reflect higher levels of staff validity

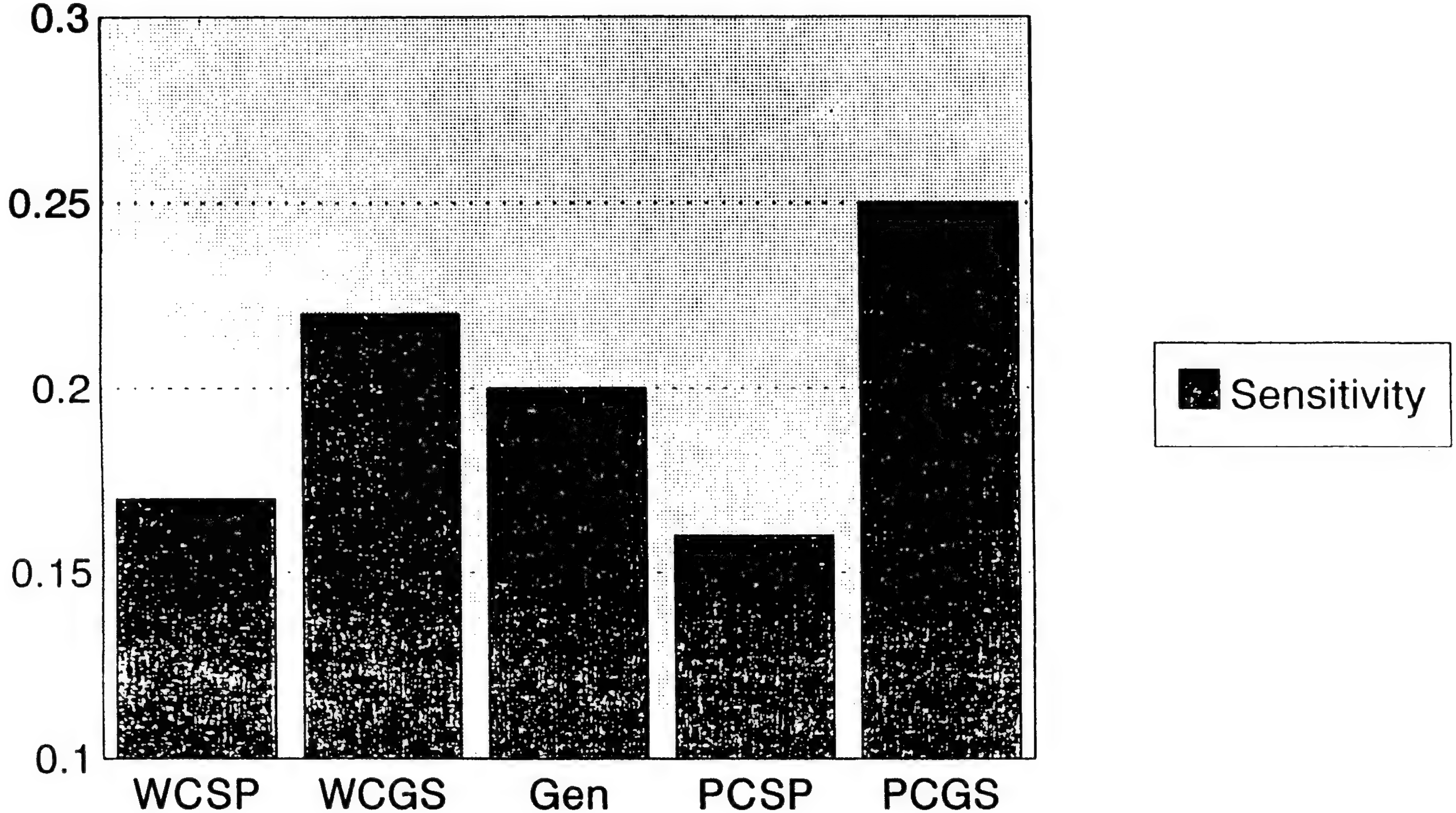
Figure 8b. Architecture and implicit coordination on staff validity

significant ($\underline{F}$ = 2.90, $\underline{df}$ = 4, 100; $\underline{p}$ < .05). The nature of these effects is presented in Figure 8c. The hierarchical sensitivity for the specialist teams was slightly better than that of the generalists, whereas the hierarchical sensitivity for the general-specialist teams was slightly worse than the generalists.

Summary of Study 1. This study replicated the three studies reported in Hollenbeck et al. (1995) and Hedlund et al. (1995), supporting the main propositions of the Multilevel Theory. The effect size reported in this study for the core constructs (i.e., .42) was well within the range established in the three previous studies (.27 and .49) and the direction of the effects was as predicted. These results also replicated the earlier findings that indicated that the core constructs mediated the effects of other variables on team decision making accuracy. That is, differences in team architecture and implicit coordination impacted team performance, but all of this influence was transmitted via the core constructs.

Overall, the results of Study 1 showed that teams of generalists and well coordinated specialists arrived at equal levels of team decision making accuracy via different routes. The well coordinated specialist teams were inferior to the generalists in terms of informity, but equal to them in terms of validity and actually superior to them in hierarchical sensitivity. The general-specialist teams, however, were disadvantaged on all three of the core constructs, and this translated into levels of decision accuracy that were highly inferior relative to that of either the generalist or well coordinated specialist teams.

Architecture, Implicit Coordination and Hierarchical Sensitivity



Lower scores reflect higher levels of hierarchical sensitivity

Figure 8c. Architecture and implicit coordination on hierarchical sensitivity

More specifically, it appeared that the reason why differences in architecture and implicit coordination affected team decision making accuracy was because of their influence on staff members' validity. Poor coordination among team members resulted in less information being available to use in decision making, and even when the information was made available, it arrived late in the trial. Hence, staff members in poorly coordinated teams also had less time to consider their judgments. In fact, an examination of the timing data revealed that the poorly coordinated general-specialists, on average, forwarded their recommendations to the leader with only 29 seconds (SD = 7) remaining in the trial, compared to 48 seconds (SD = 13) for the generalists and 41 seconds (SD = 11) for the well coordinated specialists.

This poor timing worked its way up the team hierarchy. Team leaders also had less time to consider each staff member's opinion, resulting in their developing a weighting system that was not as optimal as that found in other teams. The timing data revealed that leaders of poorly coordinated general-specialist teams, on average, had only 21 seconds (SD = 4) to consider their staff's judgments, compared to 32 seconds (SD = 9) and 29 seconds (SD = 7) for the generalists and well coordinated specialists respectively. In general, the correlation between implicit coordination and time left was −.35 ($p < .05$).

Although the effects of the core constructs as a set of main effect predictors of team decision accuracy and as mediators of the influence of noncore variables was as predicted, one aspect of the current results was somewhat surprising. In general, there was little effect for one of the core constructs--team informity--in this study. The main reason for this can be traced to the relatively high performance of the well coordinated specialist teams. These teams performed very well, despite having relatively low amounts

of total information available to the team. The important implication of this finding is that informity is a variable that might only be meaningful when making comparisons across teams that all share the same architecture. The distributed nature of the specialist teams means that each individual staff member only needs a small amount of information, and hence overall indicators of team informity are not as meaningful for such teams. For this reason, in Study 2, we describe a different operationalization of team informity for teams with highly specialized staff members.

Study 2

In general, the results from Study 1 add to the empirical base supporting the major propositions of the Multilevel Theory of team decision making. One characteristic of all the studies testing the Multilevel Theory to date, however, is that they have all studied naturally occurring variation in the core constructs. That is, the traditional paradigm for testing the theory has been to manipulate variables traditionally studied in the group decision making literature (e.g., familiarity, role redundancy, architecture) and see what impact this had on team decision making accuracy via the core constructs.

To complement the existing studies, we performed a second study for this article manipulating feedback on the core constructs, and then randomly assigning teams to conditions. This paradigm not only provides a more rigorous test of the theory, it also suggests avenues of applying the theory to improve team decision making, rather that simply using it as a vehicle to explain naturally occurring variation.

More specifically, in Study 2, we developed an on-screen decision aid that captured the team's standing on all three core constructs and provided them with graphic, up to the moment feedback on each. The feedback literature

suggests that feedback has both informational value which promotes learning and motivational value in terms of promoting effort (Ilgen, Fisher, & Taylor, 1979; Kluger & DeNisi, 1996; Taylor, Fisher, & Ilgen, 1984). Teams have been found to actively engage in feedback seeking behavior (Ancona & Caldwell, 1992) and to respond to this feedback with changes in behaviors and outcomes. Positive responses to feedback are especially likely when the feedback focuses attention on goal--performance discrepancies (Matsui, Kakayuma, & Onglatca, 1987; Mesch, Farh, & Podsakoff, 1994) and promotes learning (Kluger & DeNisi, 1996).

At the individual level, Earley, Northcraft, Lee and Lituchy (1990) demonstrated that performance information is especially useful when outcome-type feedback (expressed in terms of bottom-line results) is complemented by process feedback (expressed in terms of intervening steps that lead to bottom-line results). The increased effectiveness of outcome feedback when paired with process feedback has also been found in numerous occasions in team contexts (Anderson, Cromwell, Doman, & Howard, 1988; Flynn, Sakikibara, & Schroeder, 1995; Jones, Buerkle, Hall, Rupp, & Matt, 1993). In fact, most computerized group decision support systems (expert systems) attempt to enhance bottom-line decision making accuracy by improving decision making processes (Kleinmuntz & Schkade, 1993), although most of these are directed at individuals rather than teams (McCart & Rohrbaugh, 1989; Pinsonneault & Kraemer, 1989). When process feedback based interventions do fail (e.g., McLeod, Liker, & Lobel, 1992), it is usually because there is a weak link between the process variables and overall performance (Locke & Latham, 1994).

The Multilevel Theory would suggest that the core constructs capture the key processes that teams need to manage if their overall goal is to generate

accurate decisions, and the strong empirical relationship between these processes and overall team performance has been documented in four different studies. Therefore, we expect that teams provided with a decision aid that provides process type feedback in terms of the core constructs, will outperform teams that are only provided outcome feedback. This expectation was tested empirically in Study 2.

Method

Participants. Research participants were 380 undergraduate students at a large midwestern university who participated in 95 four-person teams. Each received course credit and was eligible to earn a cash prize based upon their team's performance.

Task. This study used the same task as Study 1--the naval command and control version of the TIDE[2]. The task was modified, however, for those research participants who were assigned to the decision aid condition (discussed more fully below). Judgments and decisions were again rendered on a seven-point continuum that varied in aggressiveness from Ignore (the lowest level of aggressiveness, to Defend (the most aggressive response).

Research design. A Solomon Four-Group design with random assignment to conditions was employed. Figure 9 depicts the application of this design in this context. Figure 10 shows how the posttest data were converted into a traditional 2 X 2 design that crossed (a) previous experience on the task (yes or no) with (b) decision aid (present or absent). Pretested groups came into the laboratory on two occasions, both of which lasted three hours. The design provided the opportunity to see if teams can immediately apply the decision aid or whether they can only apply the decision aid after having some degree of experience with the task.

Solomon Four Groups Design with Random Assignment

Condition	Time 1	Intervention	Time 2
Group 1	Pre-test	Decision Aid	Post-test 1
Group 2	Pre-test		Post-test 2
Group 3		Decision Aid	Post-test 3
Group 4			Post-test 4

Figure 9. Study 2 research design

2 X 2 Factorial Structure for Solomon Four Group Design

	Decision Aid	No Decision Aid
Pre-test	Post-test 1	Post-test 2
No Pre-test	Post-test 3	Post-test 4

Figure 10. 2 X 2 factorial structure for Solomon four group design

Measurement of team decision making outcomes. Accuracy was defined as the degree to which the team's decision (rendered by the leader) matched the correct decision given by the rules of the simulation. The performance scale and the verbal anchors associated with various outcomes were the same as in Study 1.

The core constructs were measured in the same fashion as was done in Study 1 with the exception of the team informity measure. In this study, team informity was operationalized as the total number of staff members who had all the information they needed to make a recommendation from their role. So for example, one team member was responsible for the "Motion Rule" dealing with the interaction among cues dealing with "Speed, Direction, and Angle." If, prior to the end of the trial, this person had all three of pieces off information needed to judge the aircraft from their specialized role, they were considered fully "informed." If any piece of information was missing, they were considered "uninformed." Thus, this measure could take on a value from 0 (if no staff member had all the needed information) to 3 (if all three staff members obtained what they needed).

This operationalization differs from that used in Study 1, which just looked at the total amount of information the group had on each aircraft. With highly specialized team members, the new index, which focuses directly on the specific needs of the role seemed more logical than the global, undifferentiated measure used previously. As we saw in Study 1, teams with specialized staff members can perform quite well despite not having a great deal of global, undifferentiated information.

Decision aid. The normal $TIDE^2$ computer screen presents each subject with an icon that represents the four team members engaged in the simulation.

The top of this screen contains a menu that the participants can use to measure aircraft attributes, transmit information or text messages to other team members, receive information from other team members and send recommendations to the leader (or if the person is the leader--to make decisions). There is also a count-down clock in the middle of the screen which alerts participants to how much time is left in the current trial (as well as the current trial number).

Three features were added to this screen in the present study. First, an indicator light appeared next to the icon of any team member who had obtained all the information he/she needed to make a decision based upon the team member's area of specialization. The indicator light appeared the moment that person became fully <u>informed</u>. If the light was not on, all team members knew that the person needed some additional information not available at his or her workstation.

Second, under each of the icons representing the four team members was a red bar that varied in length depending upon the <u>individual validity</u> of that staff member's recommendations (or in the case of the leader, his or her decisions). The red bar also included a numerical scale from -1.00 to +1.00. The computer simulation generated the red bar by calculating the correlation between that staff member's recommendations and the true scores for all the trials up to that moment. Like the informity indicator light, all staff members and the leader could see each others' red bar next to their icon, and hence everyone in the team knew whose recommendations were high in validity and whose were low.

Third, just under the each icon's red bar, was a green bar that varied in length depending upon the weight that the leader put on that staff member's

recommendations. This green bar also included a numerical scale from -1.00 to 1.00. The computer simulation generated the green bar by calculating the correlation between that staff member's recommendations and the leader's actual decisions for all the trials up to that moment. Like the informity indicator light and the validity bar, all staff members and the leader could see each others' green bar, and hence everyone in the team knew whose recommendations were being weighted heavily by the leader and whose recommendations were receiving less weight.

Because the red validity bar sat on top in combination with the green weight bar provided feedback on <u>dyadic sensitivity</u>. When the red bar and green bar were similar in length, this meant that the leader was weighing that staff member about as heavily as was warranted given his or her validity (high dyadic sensitivity). When the red bar exceeded the green bar, this implied that the leader was not weighing that staff member's recommendation enough, relative to his or her predictive value. When the red bar was much smaller than the green bar, this implied the leader was making the opposite mistake, that is, weighing that person's recommendation too much, relative to his or her predictive value. Thus, the larger the deviation between the two bars, the lower the dyadic sensitivity. The two bars were presented one on top of the other making comparisons easy.

Whereas staff members in teams with the decision aid worked to get their red bars as long as possible (high staff validity), the leader worked primarily to get the green and red bars for each staff member in appropriate alignment (high hierarchical sensitivity). However, as we noted in the introduction of this article, this is a dyadic process, not just a leader-driven process. Under-weighted staff members in teams with the decision aid

should work aggressively to get their leader to listen more attentively to their recommendations, while over-weighted members may keep quiet if they valued influencing their leader.

Introducing variability within staffs. In order to provide a sensitive test of the value of the decision aid, the relative importance of the different staff positions was also varied in this study. Specifically, the role of one of the staff members was weakened in terms of its potential validity by restricting the variance of the cues that made up this person's area of expertise. The lack of range in this person's cues meant there was little relationship between any recommendation coming from this station and the true score, which was for the most part, determined by variance in the cues monitored by the other two staff positions. Of course, this does not guarantee that the people occupying the other two staff positions would actually show high individual validity, but they at least had the potential to do so. The restricted cue position was highly limited in its potential in comparison to the other two. In the end, the average individual validities associated with the controlled position was .20, compared to .66 and .64 for the other positions. This puts a premium on the team's ability to arrive at an accurate and differentiated system for processing the input from the three staff members (i.e., hierarchical sensitivity).

Study 2 Results and Discussion

Table 5 shows the descriptive statistics associated with the variables measured in this study at the team level (n = 95). Analyses of these data were conducted in the same stages that typified previous studies dealing with the theory.

Table 5

Descriptive Statistics for All Variables at the Team Level (n = 95)

	Mean	SD	(1)	(2)	(3)	(4)	(5)	(6)
(1) Team Decision Accuracy	1.20	.24	---					
(2) Team Informity	2.53	.60	-.56	---				
(3) Staff Validity	.53	.11	-.70	.53	---			
(4) Hierarchical Sensitivity	.17	.09	.54	-.34	-.41	---		
(5) Experience	.50	.50	-.53	.62	.47	-.18	---	
(6) Decision Aid	.50	.50	-.25	.21	.11	-.41	-.01	---

$r > .17$; $p < .05$

Note. Decision accuracy and hierarchical sensitivity are deviation scores and hence low values reflect high standing on the constructs.

<u>The impact of the core constructs</u>. The results of the first stage of data analysis are shown in Table 6 where team decision accuracy is regressed on the three core constructs. As a set, the core constructs explained a statistically significant 63% of the predictable variance in team decision making accuracy. The bivariate correlations between the core constructs and team decision making accuracy varied in (unsigned) magnitude from a low of .54 to a high of .70. There was also a statistically significant unique effect for each of the three core constructs in the multiple regression. In terms of effect size, the majority of the unique variance was explained by team informity (incremental R^2 of .32) and staff validity (incremental R^2 of .23) relative to hierarchical sensitivity (incremental R^2 of .06) which was entered last.

As was found in three previous studies (Hedlund et al. 1996; Hollenbeck et al. 1995) but not in Study 1 of this manuscript, there was also a statistically significant two-way interaction between staff validity and hierarchical sensitivity (incremental R^2 = .02). When plotted, the nature of this interaction indicated that the combination of low hierarchical sensitivity with low staff validity led to especially poor team decision making. No other interactions were significant.

To illustrate the nature of these effects, Figure 11 shows the difference in decision accuracy and disaster rates (where a disaster is defined as a highly erroneous decision that is off by four or more points) for teams that were high on all the core constructs versus teams that were low on all the core constructs. Decision accuracy was higher for teams that were high on all three constructs by almost two standard deviation units, and the

Table 6

The Effects of the Core Constructs on Team Decision Accuracy ($n = 95$)

Hierarchical Step	Independent Variable	Total R^2	Incremental R^2
(1)	Team Informity (TI)	.32*	.32*
(2)	Staff Validity (SV)	.55*	.23*
(3)	Hierarchical Sensitivity (HS)	.61*	.06*
(4)	SV by HS	.63*	.02*

*denotes $p < .05$

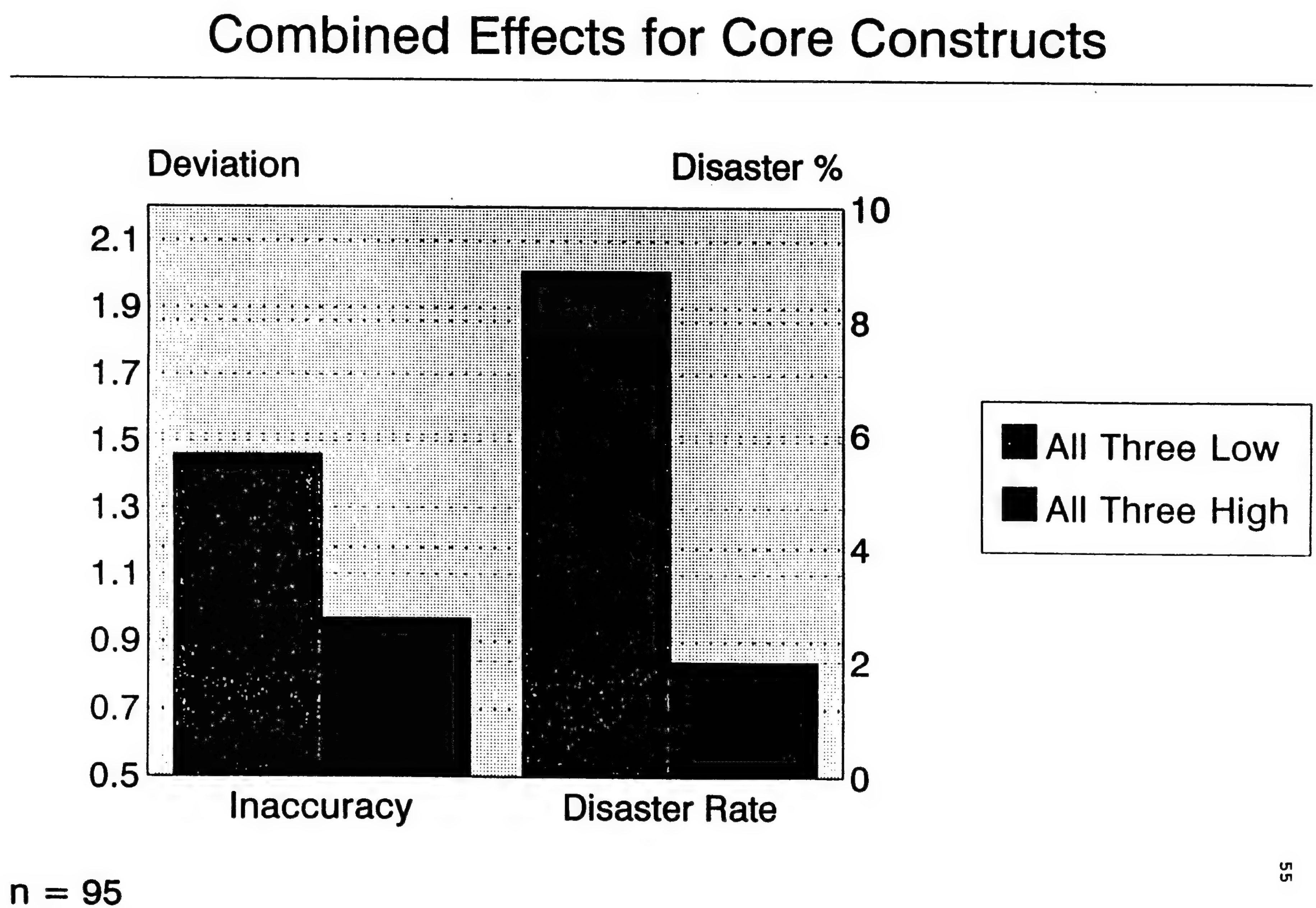


Figure 11. Combined effects for the three core constructs on team decision making accuracy and disaster rates

disaster rates for teams low on all three core constructs were four and a half times higher than the disaster rates for teams high on all three.

The impact of the noncore constructs. Table 7 shows the results for the second stage of analysis where team decision accuracy is regressed on the contrast coded variables that capture the effect of experience (i.e., being pretested), the decision aid, and the interaction between these two variables. As a set, these variables explained a statistically significant 33% of the variance in team decision making accuracy, with 27% attributable to experience and 6% attributable to the decision aid. As shown in Figure 12, as expected, experienced teams and those with the decision aid outperformed inexperienced and unaided teams. There was no interaction between experience and the decision aid, however, and thus the effect of the aid was the same for experienced and inexperienced teams.

Mediating role of the core constructs. The third stage of the analysis tested the degree to which the effects for experience and the decision aid were mediated by the core variables. Table 8 shows the results of regressing team decision accuracy on experience, the decision aid and the interaction between these two variables, after controlling for the core constructs. The results show that after entering the core constructs, the effect size for experience and the decision aid dropped from .33 to .02--a 94% reduction in the effect size. Although the effect for experience was still statistically significant (incremental R^2 of .02), the effect of the decision aid was driven to zero; decision aid effects on team decision making accuracy were fully mediated by the core constructs.

Nature of the mediating effects. To investigate the nature of the mediating effects, we treated the core constructs as dependent variables and

Table 7

The Effects of Experience and the Decision Aid on Team Decision Accuracy

Hierarchical Step	Independent Variable	Total R^2	Incremental R^2
(1)	Experience (EX)	.27*	.27*
(2)	Decision Aid (DA)	.33*	.06*
(3)	EX by DA	.33*	.00

*denotes $p < .05$

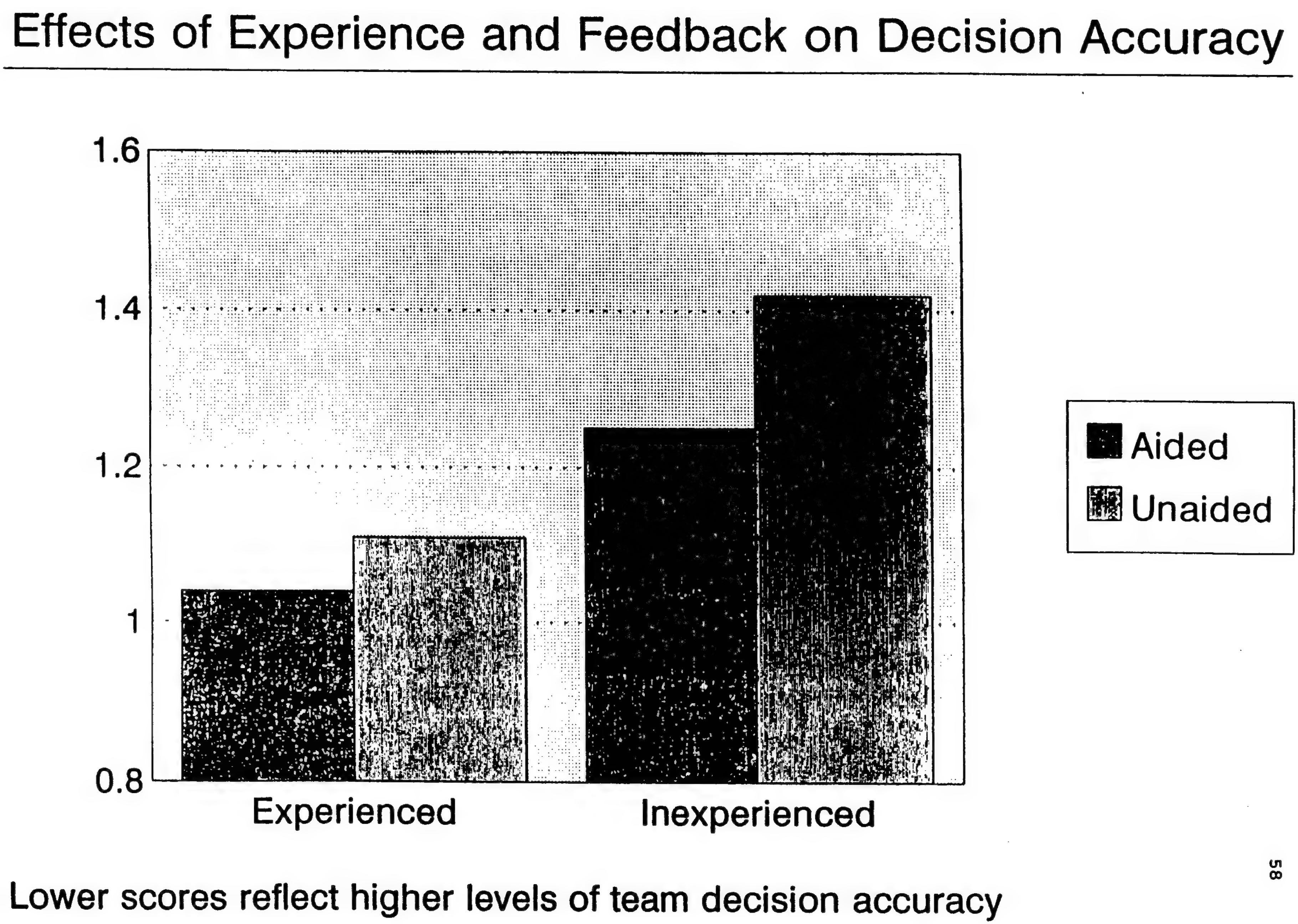


Figure 12. Effects of experience and feedback on team decision making accuracy

Table 8

The Mediating Role of the Core Constructs

Hierarchical Step	Independent Variable	Total R^2	Incremental R^2
(1)	Team Informity (TI)	.32*	.32*
(2)	Staff Validity (SV)	.55*	.23*
(3)	Hierarchical Sensitivity (HS)	.61*	.06*
(4)	SV by HS	.63*	.02*
(1)	Experience (EX)	.65*	.02*
(2)	Decision Aid (DA)	.65*	.00
(3)	EX by DA	.65*	.00

*denotes $p < .05$

examined the degree to which each was affected by experience and the decision aid. Since variability in the core constructs exists both within and across teams, repeated measures regression (Cohen & Cohen, 1983; Hollenbeck, Ilgen, & Sego, 1994) was used to decompose the nature of these effects and partition the variance.

Figures 13a, 13b and 13c show the results of this variance partitioning for each of the core constructs. For decision informity, 46% of the variance in this variable was due to between team variance (i.e., team informity), whereas 54% was attributable to variance across decisions within teams. Experience and the decision aid explained 21% of the total variance in informity, which translates into 46% (i.e., 21%/46%) of the between teams variance, which is statistically significant ($F = 25.84$; $df = 3, 91$; $p < .05$).

The main effects for both experience and the decision aid were statistically significant. As shown in Figure 14a, the nature of these effects indicated, as expected, greater experience and presence of the decision aid led to greater levels of informity. In terms of magnitude, the effect size for experience (incremental R^2 of .39) was larger than that for the decision aid (incremental R^2 of .04). There was also a marginally significant interaction (incremental $R^2 = .02$; $F = 3.02$; $p = .08$) that suggested the aid had more impact on inexperienced teams relative to those with experience on the task.

In testing the effects for individual validity, one needs to recall that the validity of one of the staff members was manipulated so as to be close to zero and hence could not vary. Thus, the effects of the noncore constructs on individual validity were only tested on the staff members who could actually alter their validity through increased effort or altered task strategies.

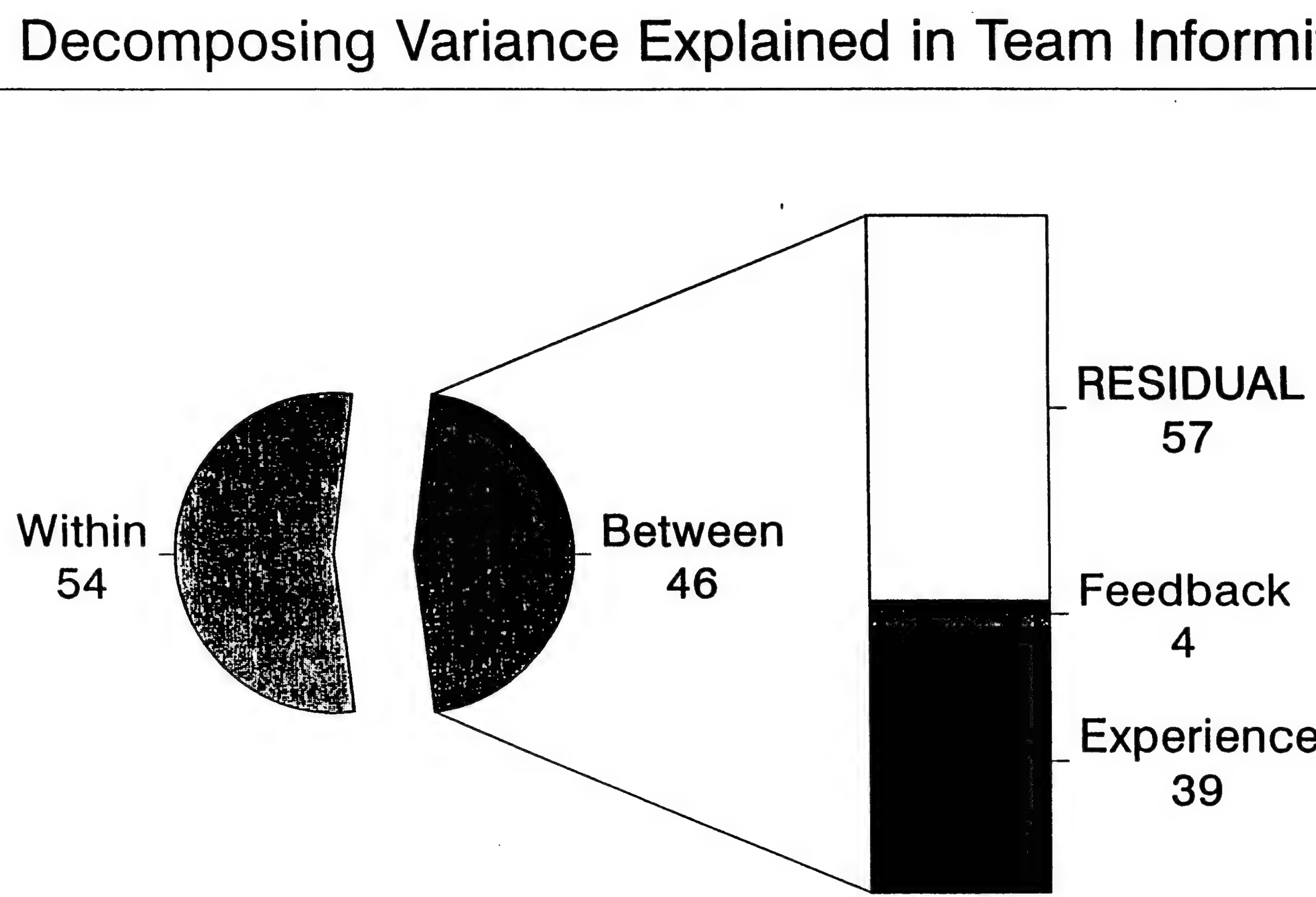


Figure 13a. Decomposing variance explained in team informity

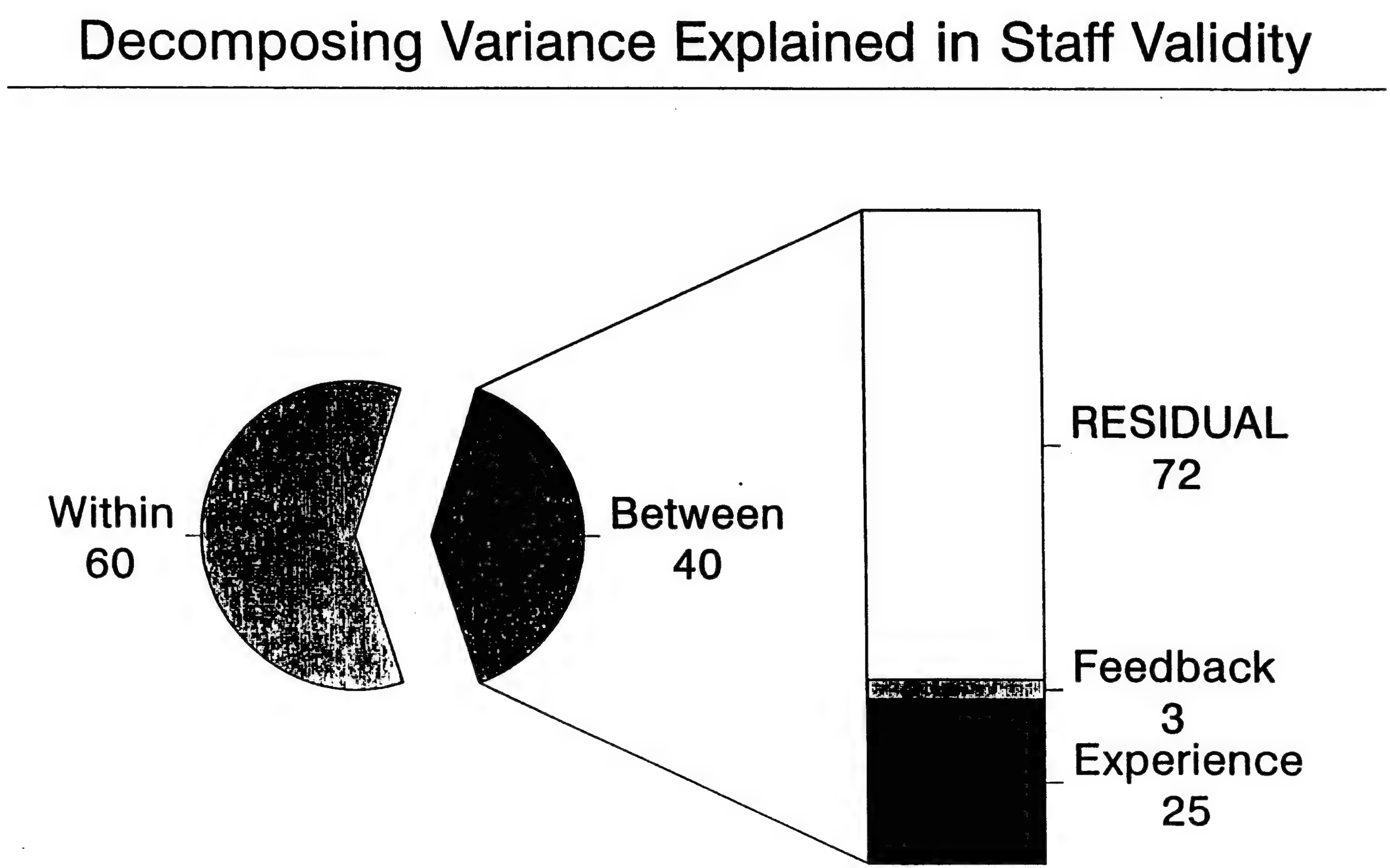


Figure 13b. Decomposing variance explained in staff validity

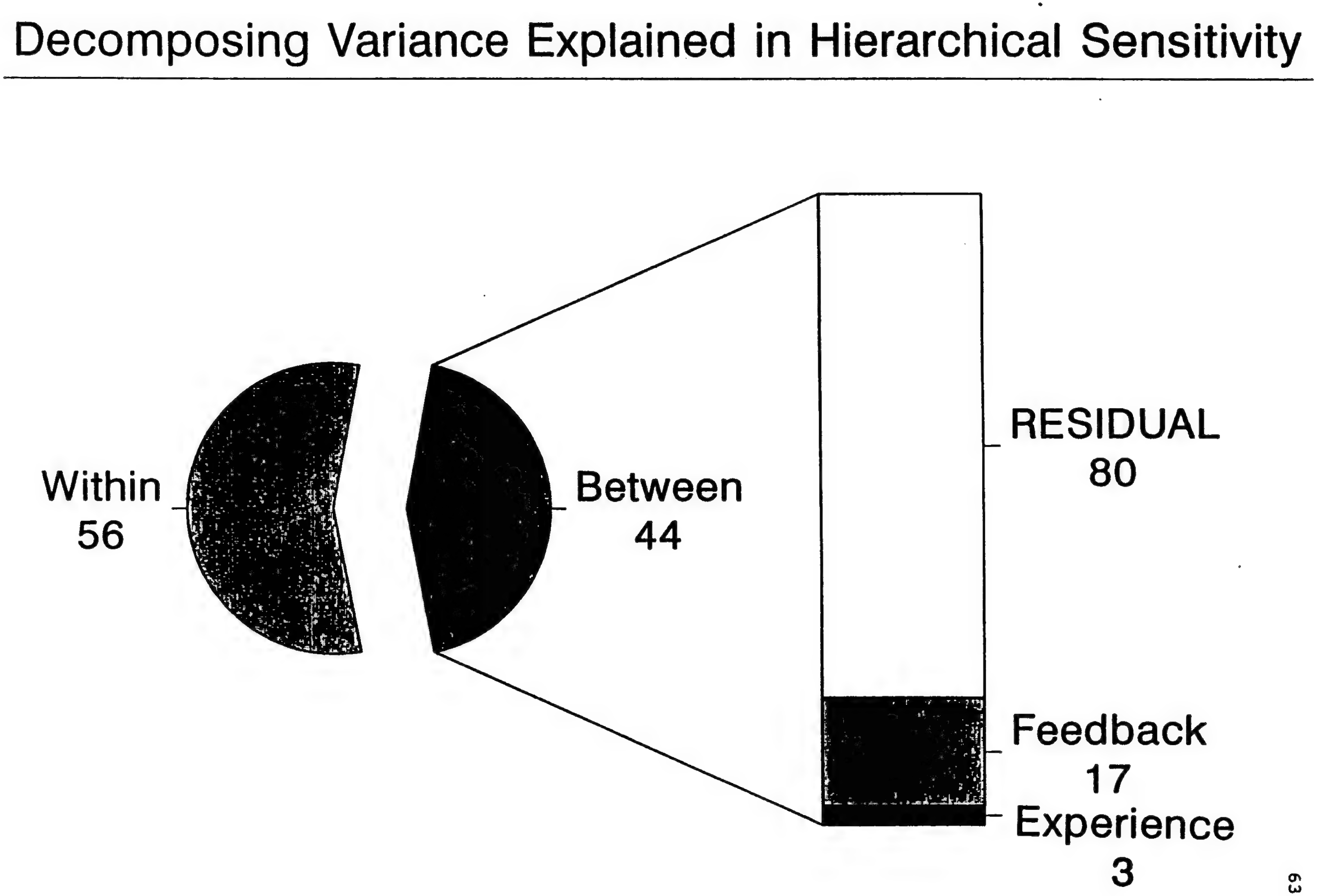


Figure 13c. Decomposing variance explained in hierarchical sensitivity

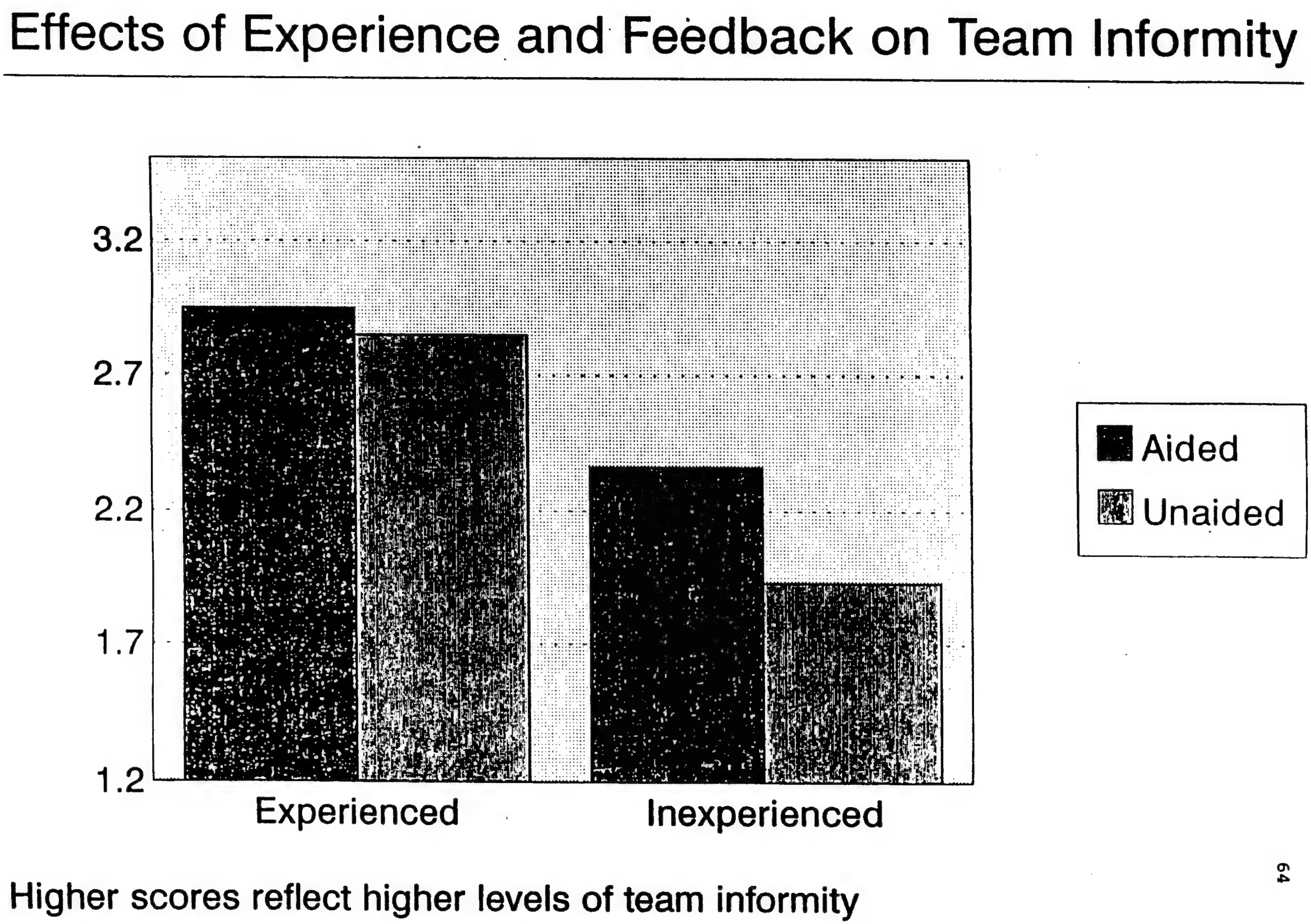


Figure 14a. Effects of experience and feedback on team informity

Thus, the number of observations for this analysis is 190 (two staff members in each of 95 teams) not 285.

For individual validity 60% of the total variance in this variable was due to between team variance (i.e., staff validity), whereas 40% was attributable to variance across staff members within teams. Experience and the decision aid explained 17% of the total variance in staff validity, which translates into 28% (i.e., 17%/60%) of the between teams variance. This was statistically significant (F = 11.79, df = 3, 91; $p < .05$). There were statistically significant main effects in the predicted direction for both experience and the decision aid, although as shown in Figure 14b, the magnitude of the effect for experience (incremental R^2 of .25) was larger than that for the aid (incremental R^2 of .03).

For dyadic sensitivity, 44% of the total variance was due to between team variance, whereas 56% was attributable to variance across the three leader-staff dyads within teams. Experience and the decision aid explained 9% of the total variance in hierarchical sensitivity, which translates into 20% (i.e., 9%/44%) of the between teams variance, which was statistically significant (F = 7.58, df = 3, 91; $p < .05$). Of the three variables, only the main effect for the decision aid was statistically significant. As predicted, the presence of the aid led to enhanced sensitivity, explaining 17% of the variance, and the nature of these results are shown in Figure 14c.

Summary of Study 2. The study replicated several aspects of previous studies testing the Multilevel Theory. The core constructs explained a large percentage of variance in overall team decision making accuracy (61%) and the direction of the effects was as predicted. These results also replicated the mediation role of the core constructs with respect to other variables as they

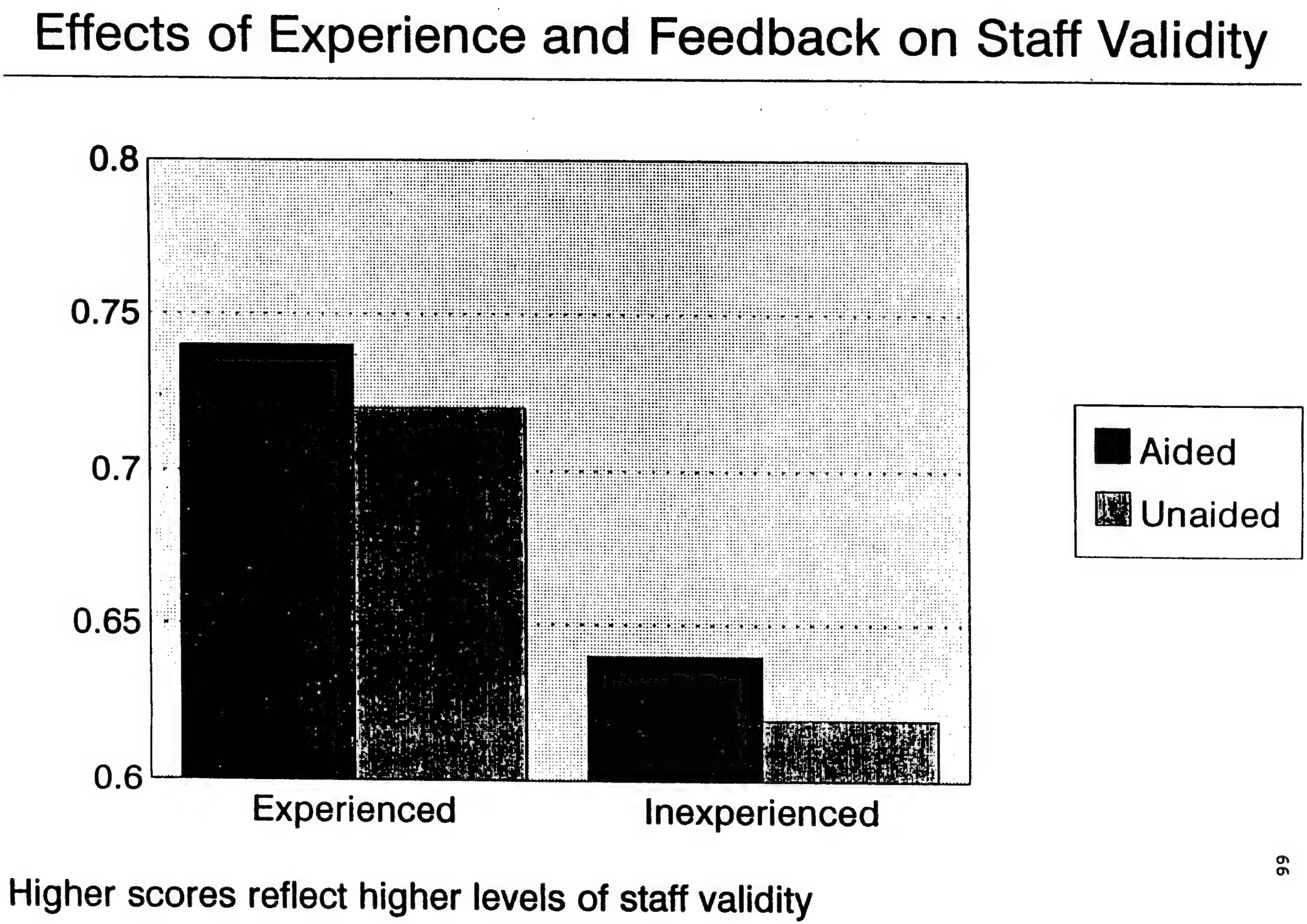
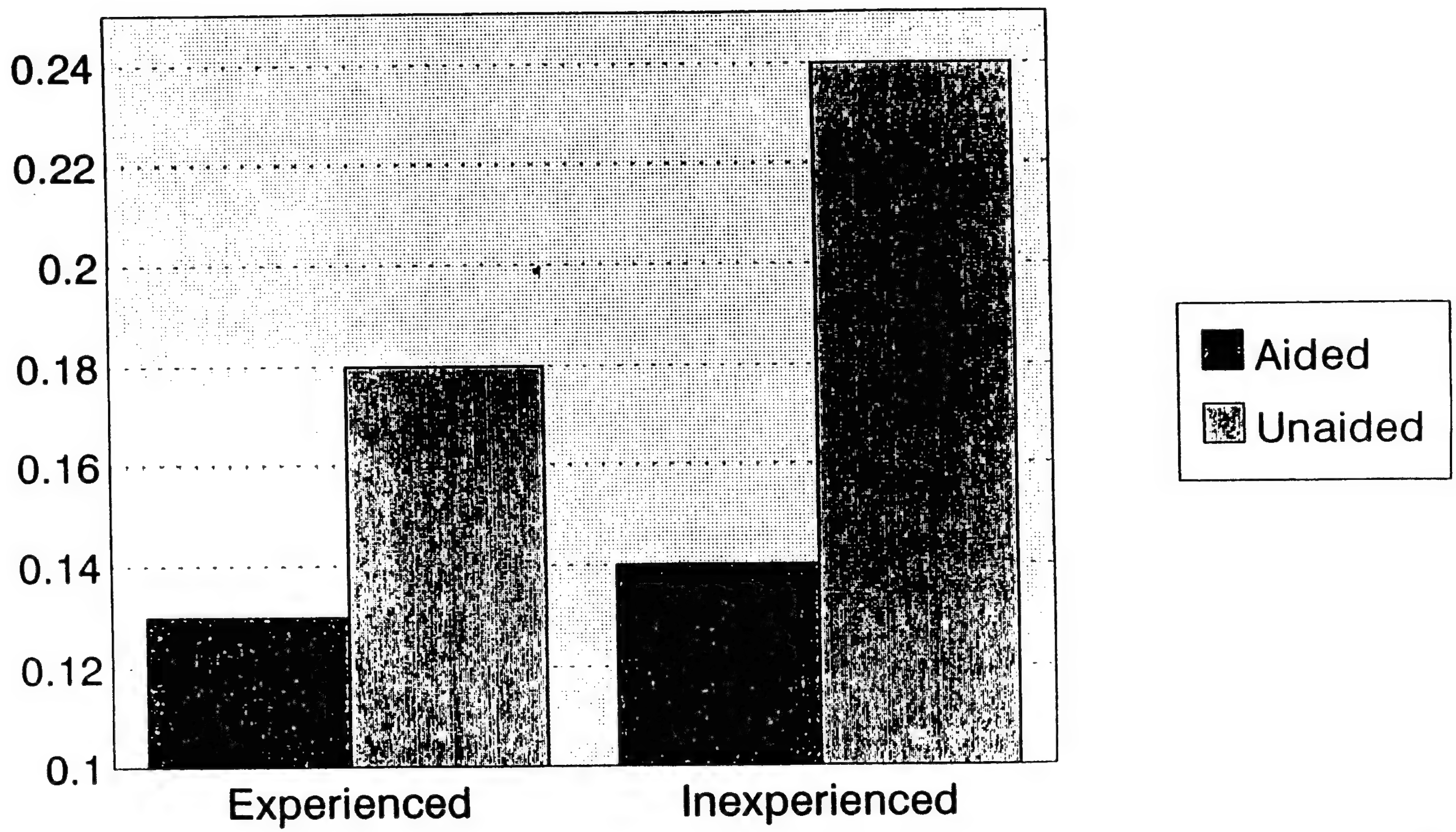


Figure 14b. Effects of experience and feedback on staff validity

Effects of Experience and Feedback on Hierarchical Sensitivity



Lower scores reflect higher levels of hierarchical sensitivity

Figure 14c. Effects of experience and feedback on hierarchical sensitivity

affected team decision making accuracy. Over 90% of the effects for experience and feedback were explained by the core constructs. Indeed, all of the variance attributable to the feedback intervention aid was accounted for by the core constructs.

Overall, the results of Study 2 showed that experienced teams and those with the decision aid outperformed inexperienced and unaided teams. The main virtue of the decision aid was that it promoted hierarchical sensitivity. That is, presence-absence of the decision aid had a major impact on hierarchical sensitivity ($r = -.41$) which in turn had a major impact on overall team decision making accuracy ($r = .50$). The decision aid also had beneficial effects on team informity and staff validity, but these effects were smaller in magnitude compared to the effects on hierarchical sensitivity. Instead, these variables were affected mainly by experience. That is, experience led to higher team informity ($r = .62$) and staff validity ($r = .47$) which in turn led to higher overall team performance ($r = -.56$ and $-.70$ respectively).

Discussion

Overall Findings on the Core Constructs

The findings of both Study 1 and Study 2 replicated earlier research testing the Multilevel Theory. The results indicate that the core constructs specified by this theory explain a great deal of the variance in team decision making accuracy. The core constructs also mediated most of the effects of group architecture and implicit coordination in Study 1 and experience and the on-screen decision aid developed in Study 2. In fact, since the measure of implicit coordination could also be constructed for Study 2, we attempted to replicate Study 1's finding with Study 2 data. We found that implicit

coordination explained 16% of the variance in team decision making accuracy in Study 2, and all of this was again mediated by the core constructs.

If one adds the two studies reported here in with the three studies reported earlier, the scope and consistency of the empirical base associated with this theory becomes clearer. Taken in total, the five studies reported to date have involved 386 teams comprised of 1,544 individuals who have made 17,110 team decisions and 68,440 staff recommendations. The average amount of variance in team decision making accuracy accounted for by the core constructs is 45% (range = .27 to .63) and, on average, 89% of the variance attributable to noncore constructs has been explained by the core constructs (range = 72% to 100%). In addition, in four out of five cases, there has been a statistically significant interaction between hierarchical sensitivity and staff validity.

Variability in Effects Between Core Constructs

Although the effects for the three core constructs as a set have been consistent throughout the five studies published or conducted during the time period supported by this project, the evidence suggests the isolated impact of each core construct differs in terms of both magnitude and variability across studies. Examination of these differences reveals a great deal about the role of each in different types of decision making contexts.

For example, the bivariate effect for team informity is the most variable of the three core constructs across the five studies, ranging from a low of .19 to a high of .62. Much of this variability can be traced to the degree of specialization built into the staff positions and the means of operationalizing team informity. As we saw in Study 1, with specialized staff members, the total amount of information held by the team is not as important

as the configuration of information distributed across staff members. In Study 1, the well coordinated specialists had a great deal less information regarding the aircraft as a whole than the generalist teams, however, they performed just as well (weakening the bivariate correlation between these two variables).

Based upon this finding, in Study 2, we operationalized team informity more narrowly, in terms of how many staff members had all the information they needed to make a recommendation based upon their specialty. This more focused operationalization had a bivariate correlation of .57 with overall team decision making accuracy. The total information operationalization would have generated a much lower correlation in this context ($r = .17$). Moreover, the on-screen decision aid, which was targeted at getting the right distribution of information rather than maximizing the total amount of information for each staff member had a statistically significant impact on the narrowly focused measure of team informity, but not the more global measure. Thus, unaided teams had the same amount of information as the aided teams, they just did not have the right information distributed to the right person. Thus, with teams of distributed specialists, "more is not necessarily better" in terms of information. The configuration of information and the efficiency with which this information gets distributed seems more important than the sheer amount of information.

Whereas the effect for team informity manifests the most variability across studies, the effect for hierarchical sensitivity seems to be the lowest in magnitude across the studies (average $r = .31$ versus .40 for team informity and .52 for staff validity). Part of this is probably due to unreliability in the index, which is a function of the number of decisions upon which the index

is based. Since the measure of dyadic sensitivity is derived from a multiple regression of the true score and the leader's decision of the staff's recommendations, the stability of this measure depends a great deal on the number of decisions used in the regressions. The effects for this variable tends to be weak (e.g., r's between .12 and .18) when the number of decisions is small (e.g., in the 22 to 24 range as in Hollenbeck et al. 1995, Study 2 and Hedlund et al., 1996), but stronger when the number of decisions is greater than 40 (r's between .27 and .54 in the remaining studies).

In addition to methodological issues, however, hierarchical sensitivity is probably the most complex of the core constructs and the most difficult one for the team to learn. Without any aid, the leader and staff members probably have only a general idea of the relative value of each staff member or the weight being assigned to each staff member during real-time task engagement. This limits the teams ability to develop a highly accurate and differentiated weighting structure. For example, in both studies reported here, the weight the leader placed on staff recommendations was too high (average weight = .33 and .37 in Study 1 and 2 respectively) relative to their ideal weight (average = .23 and .33 respectively). A similar picture emerges if one looks at the overall amount of variance in the leader's decision accounted for by the staff members' recommendations versus the amount of variation explained by the staff recommendations in the true decision. In both studies, the staff's recommendations explained roughly twice as much variance in the leader's decision relative to the true decisions. The leaders also tended to show less variability in their weighting of staff members (SD = .06 and .16 for Study 1 and Study 2 respectively) relative to the variability associated with the ideal set of weights (SD = .10 and .28). Thus, the leaders relied too heavily

on the staff and failed to sufficiently differentiate among staff members relative to the ideal.

Given the complexity of the hierarchical sensitivity construct, it is instructive to break it down into its constituent parts to get a better grasp of why the unaided teams performed so poorly on this factor relative to the teams that had the decision aid. To do this, we reran the regression reported in Figure 13c, using Edwards' (1995) regression decomposition techniques to see exactly how the decision aid affected hierarchical sensitivity. With Edwards' technique, one simply follows up a regression where a simple difference score is the dependent variable, with two additional regressions treating each component of the difference score as the criterion. The overall effect for the difference score is then decomposed into two component parts.

In decomposing these regressions we found that, although both experience and the intervention affected hierarchical sensitivity (i.e., the difference score), neither of these two factors affected either one of the components that went into hierarchical sensitivity. That is, having the decision aid did not cause leaders to raise or lower their overall weight they assigned to staff member judgments, nor did it affect the ideal weights for staff members. Rather, the effect of the decision aid on hierarchical sensitivity was due to its generating a more individually tailored set of leader weights for staff members.

Relationships Among the Core Constructs

Another fact made clear by examining the results of all five studies conducted to date on the Multilevel Theory is that there are strong, probably causal, relationships among the core constructs. Although the initial development of the theory (see Hollenbeck et al., 1995, pp. 296-298) described

the core constructs as emanating from lower levels to higher levels (e.g., informity at the decision level fed into validity at the individual level which fed into sensitivity at the dyadic level) no explicit relationships among the core constructs were specified originally. However, it is certainly reasonable to expect that the better informed the staff is, the higher their validity is likely to be. Further, the dyadic nature of hierarchical sensitivity implies that having a highly knowledgeable staff should promote sensitivity. A knowledgeable staff may be more likely to "know what they don't know," and hence warn the leader when they have low confidence in their recommendations, thus promoting the leader's sensitivity.

Empirically, all five studies to date have documented significant relationships between core constructs residing at adjacent levels. For example, the average correlation between team informity and staff validity across the five studies is .39 (with a range of .24 to .64). Similarly, the average correlation between staff validity and hierarchical sensitivity across studies is -.31 (range -.12 to -.41). Thus, whereas the theory focuses primarily on the relationships between the non-core variable and the core variables, the data suggests that propositions regarding the relationships among the core constructs should also be included.

<u>Moderators of the team informity-staff validity relationship</u>. To further our knowledge of the relationships between the core constructs in Study 2, we tested several possible factors that might affect these relationships. For example, under what conditions is being informed or uninformed likely to have the most or least impact on staff validity? The literature on decision making suggests that under stress, individuals have a more difficult time processing large amounts of information (Janis & Mann, 1977; Tjosvold, 1987). Thus, one

might speculate that being well-informed under high time pressure is less of an advantage than being well-informed under low time pressure.

We tested this speculation with Study 2 data by operationalizing stress as the amount of time left when the staff entered their final recommendations. That is, there was a time deadline associated with each decision, and the computers that the teams were working on emanated a recurring warning sound when there was less than 30 seconds to the deadline. Earlier construct validation research with this simulation (Hollenbeck, Sego, Ilgen, Major, Hedlund, & Phillips, in press) showed that physiological reactions (heart rate, pulse, and blood pressure) increased for participants as the deadline neared.

To see if this type of stress might affect the team informity-staff validity relationship, we regressed staff validity on team informity, time left when the recommendations were made, and the interaction between these variables. Since the validity of one of the staff positions was fixed, we focused primarily on the staff validity of the other two positions.

The results indicated that team informity explained 36% of the variance in staff validity under these circumstances, and while there was no main effect for "time left," the interaction between time left and informity was significant (an incremental R^2 of .04; F = 6.49, p <.05). As predicted, the nature of this interaction was such that the high levels of informity were more useful when the teams were not under time pressure relative to when they were under time pressure.

The literature on social loafing would also suggest that accountability is one of the key factors in ensuring that each team member contributes to the best of his or her ability (Miles & Greenberg, 1993; Williams, Harkins, &

Latane, 1981; Williams, Nida, Baca, & Latane, 1989). Because the decision aid displayed the value of each staff member's contribution so publically, we expected that accountability would be higher for staff members working with the aid, relative to staff members where no such graphic information regarding each person's contribution was available. Moderated regression analyses supported this prediction. The interaction between team informity and the decision aid accounted for a statistically significant increment in R^2 over and above that attributable to team informity (incremental $R^2 = .05$, $F = 6.39$, $p < .05$). The nature of this interaction suggested that high levels of informity were more likely to translate into high levels of validity when staff members' contributions were made public via the decision aid than when they were less readily apparent.

<u>Moderators of the staff validity-hierarchical sensitivity relationship</u>. We also sought to find the conditions where staff validity to affect hierarchical sensitivity. Using much the same logic as with team informity, we first tested the degree to which the this relationship would be attenuated by time pressure. However, rather than using the time left when the staff made their recommendations, we instead looked at the time pressure experienced by the leader (operationalized as time left when he or she made the final team decision). That is, the leader would probably have a more difficult time making fine discriminations among the varied staff members when under high time pressure relative to low time pressure.

The moderated regression results did not support this speculation. Staff validity explained 17% of the variance in hierarchical sensitivity, but there was no main effect for time pressure experienced by the leader nor was the

interaction between this variable and staff validity statistically significant.

As we noted earlier, given the complex nature of hierarchical sensitivity, this is probably the most difficult of the core constructs for the team to manage and learn. The general tendency is for leaders to weigh staff recommendations too heavily and too uniformly. This basic tendency has to be unlearned via experience. Assuming this would take time, we predicted that the negative effect of having a poor staff member on hierarchical sensitivity would be especially great when the leader lacked experience. To test this we repeated the procedure listed above substituting experience as the potential moderator. Although experience had no main effect on hierarchical sensitivity, the interaction between experience and staff validity was marginally significant, accounting for an increment of 3% variance explained in this outcome ($F = 3.46$, $p = .07$). The nature of this interaction was in the predicted direction, in that the deleterious effect of having a poor staff was especially strong when the team lacked experience.

Thus, taken as a whole, these additional analyses not only reveal strong relationships among the core constructs, but suggest that the strength of these relationships varies under different conditions. For example, although the decision aid had no direct impact on staff validity, it did enhance the relationship between team informity and staff validity. Similarly, whereas experience had no direct impact on hierarchical sensitivity, it did enhance the relationship between staff member validity and hierarchical sensitivity. Time pressure also affected one of these relationships, in the sense that high time pressure experienced by the staffs weakened the relationship between team informity and staff validity.

Conclusion

This research program supported by the Office of Naval Research has accomplished four primary objectives. First, it led to the development of a research paradigm for studying team decision making in hierarchical teams with distributed expertise. A simulation known by its acronym, $TIDE^2$, was developed and laboratory space with a number of personal computers (15) on a local network was dedicated to team research. This simulation has proven very valuable for data collection on teams where the teams are to be presented with multi-attribute decision objects and asked to reach a series of decisions over time under controlled conditions. The task has been used in a number of studies, many of which have been published in professional journals and presented at conferences (see Appendix A). The task is also being used in other team laboratories.

Second a theory of team decision making was developed. Research that has been conducted using the theory (including that described in this report) has been quite supportive of the importance of its three core constructs with respect to their influence on the accuracy of team decisions. Across a variety of studies, from 30-50% of the predictable variance in team decision making accuracy is explained by the three constructs. The importance of these constructs has been demonstrated by the fact that they mediate the effect of conditions in task and social environments of teams. In some studies, the mediation was only partial, but in a sizable majority of cases, the mediation was complete. Thus, importance of the constructs and the fact that they act as mediators has been well established in this research, although it must be kept in mind that the setting for the research has tended to be limited to the $TIDE^2$ simulation.

Third, the research has looked at a number of important conditions faced by decision making teams and found the research paradigm and the theory useful for looking at the effects of a large number of conditions on team decision making accuracy. These conditions are the ones that have been investigated for some time, for example, the medium through which team members communicate, the time pressure under which decisions are made, the ambiguity of information used to reach a decision, social-emotional factors, such as the degree of familiarity among team members, and the way in which team interactions are structured. Such factors are typically studied ahistorically by observing the effects of these conditions on some critical team characteristics. We have done the same.

Finally, the most recent work has investigated team processes dynamically looking at team performance over time. The study that provided feedback geared to the three core constructs that allowed teams to learn by improving their informity, staff validity and hierarchical sensitivity both explored teams' ability to learn from their past behavior and the usefulness of an intervention that focused on the core constructs of the model. The latter results were encouraging. They suggest that future research should explore how teams learn from their past behavior then adapt and adjust to future needs on the basis of what they know. The results of the past research suggests that the model used here may be useful for guiding research on adaptability and change in decision making teams with distributed expertise.

References

Ancona, D. G., & Caldwell, D. F. (1992). Bridging the boundary: External activity and performance in organizational teams. Administrative Science Quarterly, 37, 634-665.

Anderson, D. C., Cromwell, C. R., Doman, M., & Howard, G. S. (1988). Performance posting, goal setting, and activity contingent praise applied to a university hockey team. Journal of Applied Psychology, 73, 87-95.

Brehmer, B., & Hagafors, R. (1986). Use of experts in complex decision making: A paradigm for the study of staff work. Organizational Behavior and Human Decision Processes, 38, 181-195.

Carley, K. M. (1991). Designing organizational structures to cope with communication breakdowns: A simulation model. Industrial Crisis Quarterly, 5, 19-57.

Carley, K. M. (1992). Organizational Learning and personnel turnover. Organization Science, 3, 20-46.

Carley, K. M., & Lin, Z. (1995). Organizational designs suited to high performance under stress. IEEE Transactions on Systems, Man, and Cybernetics, 25, 221-230.

Cannon-Bowers, J. A., Salas, E., & Converse, S. A. (1990). Cognitive psychology and team training: Shared mental models in complex systems. Human Factors Bulletin, 33, 1-4.

Cohen, J., & Cohen, P. (1983). Applied multiple regression/correlation analysis for the behavioral sciences. Hillsdale, NJ: Erlbaum.

Earley, P. C., Northcraft, G. B., Lee, C., & Lituchy, T. R. (1990). Impact of process and outcome feedback on the relation of goal setting to task performance. Academy of Management Journal, 33, 87-105.

Edwards, J. R. (1995). Alternatives to difference scores as dependent variables in the study of congruence in organizational research. Organizational Behavior and Human Decision Processes, 64, 307-324.

Flynn, B. B., Sakikibara, S., & Schroeder, R. G. (1995). Relationship between JIT and TQM: Practices and performance. Academy of Management Journal, 38, 1325-1360.

Hedlund, J., Ilgen, D. R., & Hollenbeck, J. R. (1996). The effect of computer-mediated versus face-to-face communication on decision making in hierarchical teams. Unpublished manuscript. East Lansing: Michigan State University.

Hollenbeck, J. R., Ilgen, D. R. Sego, D., Hedlund, J., Major, D. A., & Phillips, J. (1995). The multilevel theory of team decision-making: Decision performance in teams incorporating distributed expertise. Journal of Applied Psychology, 80, 292-316.

Hollenbeck, J. R., Ilgen, D. R., & Sego, D. J. (1994). Repeated measures regression: Enhancing the power of leadership research. Leadership Quarterly, 5, 3-23.

Hollenbeck, J. R., Sego, D. J., Ilgen, D. R., Major, D. A., Hedlund, J., & Phillips, J. (in press). Team decision making accuracy under difficult conditions: Construct validation of potential manipulations and measures using the TIDE^2 simulation. In M. Brannick, E. Salas, & C. Prince (Eds.), New directions in team measurement. San Francisco: Jossey-Bass.

Ilgen, D. R., Fisher, C. D., & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organizations. Journal of Applied Psychology, 64, 349-371.

Ilgen, D. R., Major, D., Hollenbeck, J. R., & Sego, D. (1993). Team research in the 90's. In M. Chemers & R. Aymon (Eds.), Leadership and organizational effectiveness (pp. 245-270). New York: Academic Press.

Janis, I. L., & Mann, L. (1977). Decision making. New York: Free Press.

Jones, S. D., Buerkle, M., Hall, A., Rupp, L., & Matt, G. (1993). Work group performance measurement and feedback: An integrated comprehensive system for a manufacturing department. Group and Organization Management, 18, 269-291.

Kleinmuntz, D. N., & Schkade, D. A. (1993). Information displays and decision processes. Psychological Science, 4, 221-227.

Klimoski, R., & Mohammed, S. (1994). Team mental model: Construct or metaphor? Journal of Management, 4, 403-439.

Kluger, A., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. Psychological Bulletin, 119, 254-284.

Locke, E. A., & Latham, G. P (1994). Comments on McLeod, Liker, and Lobel. Journal of Applied Behavioral Science, 28, 42-45.

Matsui, T., Kakayuma, T., & Onglatca, M. L. (1987). Effects of goals and feedback on performance in groups. Journal of Applied Psychology, 72, 407-415.

Mesch, D. J., Farh, J., & Podsakoff, P. M. (1994). Effects of feedback sign on group goal setting, strategies and performance. Group and Organization Management, 19, 309-333.

McCart, A. T., & Rohrbaugh, J. (1989). Evaluating group decision support systems effectiveness: A performance study of decision conferencing. Decision Support Systems, 5, 243-253.

McLeod, P., Liker, J. K., & Lobel, S. A. (1992). Journal of Applied Behavioral Science, 28, 15-41.

Miles, J. A., & Greenberg, J. (1993). Using punishment threats to attenuate social loafing effects among swimmers. Organizational Behavior and Human Decision Processes, 56, 246-265.

Orasanu, J. (1990). Shared mental models and crew decision making (Tech. Rep. #46). Princeton, NJ: Princeton University, Cognitive Sciences Laboratory.

Pinsonneault, A., & Kraemer, K. L. (1989). The impact of technological support on groups. Decision Support Systems, 5, 197-216.

Rogelberg, S. G., Barnes-Farrell, J. L., & Lowe, C. A. (1992). The step-ladder technique: An alternative group structure facilitating effective group decision making. Journal of Applied Psychology, 77, 730-737.

Rouse, W. B., Cannon-Bowers, J. A., & Salas, E. (1992). The role of mental models in team performance in complex systems. IEEE Transactions on Systems, Man, and Cybernetics, 22, 1296-1308.

Sniezek, J. A., & Buckley, T. (1995). Cueing and cognitive conflict in judge-advisory decision making. Organizational Behavior and Human Decision Processes, 62, 159-174.

Taylor, M. S., Fisher, C. D., & Ilgen, D. R. (1984). Individuals reaction to feedback in organizations: A control theory perspective. In K. M. Rowland & G. R. Ferris (Eds.), Research in personnel and human resource management (pp. 81-124). Greenwich, CT: JAI Press.

Tjosvold, D. (1982). Effects of approach to controversy on superiors' incorporation of subordinates' information in decision making. Journal of Applied Psychology, 67, 189-193.

Tjosvold, D. (1987). Flight crew collaboration to manage safety risks. Group and Organization Studies, 15, 177-191.

Williams, K. D., Harkins, S. G., & Latane, B. (1981). Identifiability as a deterrent to social loafing: Two cheering experiments. Journal of Personality and Social Psychology, 40, 303-311.

Williams, K. D., Nida, S. A., Baca, L. D., & Latane, B. (1989). Social loafing and swimming: Effects of identifiability on individual and relay performance of intercollegiate swimmers. Basic and Applied Social Psychology, 10, 73-81.

APPENDIX A

Products Produced on
Office of Naval Research
Team Decision Making in Hierarchical Teams
with Distributed Expertise

N00014-90-J-1786(1 April 1990-31 August 1993)
N00014-93-0983 and N00014-93-0983 mod. no. A00001
(1 August 1993 through 31 October 1996)
N00014-93-1-1385(15 September 1993 through 14 September 1996)

D. R. Ilgen & J. R. Hollenbeck, PIs
Michigan State University

1 April 1990- 15 December 1996

Journal Articles (Including those under review)

1. Hollenbeck, J. R., Ilgen, D. R., Phillips, J., & Hedlund, J. (1994). Decision risk in dynamic two-stage contexts: Beyond the status quo. Journal of Applied Psychology, 79, 592-598.

2. Hollenbeck, J. R., Ilgen, D. R., & Sego, D. J. (1994). Data analytic methods for studying team phenomena: A case for repeated measures designs. Invited paper for special issue of Leadership Quarterly, 5, 3-23.

3. Hollenbeck, J. R., Ilgen, D. R., Sego, D. J., Hedlund, J., Major, D. A., & Phillips, J. (1995). The multi-level theory of team decision making: Decision performance in teams incorporating distributed expertise. Journal of Applied Psychology, 80, 292-316.

4. Quinones, M. A. (1995). Pretraining context effects: Training assignment as feedback. Journal of Applied Psychology, 80, 226-238.

5. Hollenbeck, J. R., Ilgen, D. R., Tuttle, D., & Sego, D. J. (1995). Team performance on monitoring tasks: An examination of decision errors in contexts requiring sustained attention. Journal of Applied Psychology, 80, 685-696.

6. Phillips, J. M., Hollenbeck, J. R., & Ilgen, D. R. (1996). The prevalence and prediction of positive discrepancy creation: An integration of episodic and non-episodic theories of motivation. Journal of Applied Psychology, 81, 498-511.

Technical Reports

7. Hollenbeck, J. R., Sego, D. J., Ilgen, D. R., & Major, D. A. (1991). Team interactive decision exercise for teams incorporating distributed expertise (TIDE^2): A program and paradigm for team research (Tech. Rep. 91-1). East Lansing: Michigan State University.

8. Ilgen, D. R., Major, D. A., Hollenbeck, J. R., Sego, D. J. (1991). Decision making in teams: Raising an individual decision making model to the team level (Tech. Rep. 91-2). East Lansing: Michigan State University.

9. Barrett, L. E. (1992). Decision making teams: Their study in the U. S. military (Tech. Rep. 93-1). East Lansing: Michigan State University.

10. Ilgen, D. R., & Hollenbeck, J. R. (1993). Effective team performance under stress and normal conditions: An experimental paradigm, theory and data for studying team decision making in hierarchical teams with distributed expertise (Final Rep. NR. 93.-2). East Lansing: Michigan State University.

Book Chapters

11. Ilgen, D. R., Major, D. A., Hollenbeck, J. R., & Sego, D. J. (1993). Team research in the 1990s. In M. M. Chemers & R. Ayman (Eds.), Leadership theory and research: Perspectives and directions (pp. 245-270). New York: Academic Press.

12. Ilgen, D. R., Major, D. A., Hollenbeck, J. R., & Sego, D. J. (1995). Raising an individual decision-making model to the team level: A new research model and paradigm. In R. Guzzo & E. Salas (Eds.), Team decision making in organizations (pp. 113-148). San Francisco: Jossey-Bass.

13. Hollenbeck, J. R., LePine, J., & Ilgen, D. R. (in press). Individual differences in adapting to roles in decision making teams: Effects on assimilation of new members into interacting teams. In K. R. Murphy (Ed.), Individual differences in organizations (pp. 300-323). San Francisco: Jossey-Bass.

14. Hollenbeck, J. R., Sego, D. J., Ilgen, D. R., Major, D. A., Hedlund, J., & Phillips, J. (1994). Team decision making accuracy under difficult conditions: Construct validation of potential manipulations and measures using the TIDE^2 simulation. In M. Brannick, E. Salas, & C. Prince (Eds.), New directions in team measurement. San Francisco: Jossey-Bass.

15. Ilgen, D. R., LePine, J. A., & Hollenbeck, J. R. (in press). Effective decision making in multinational teams. In P. C. Earley & M. Erez (Eds.), New perspectives in international industrial and organizational psychology. San Francisco: Jossey-Bass.

16. Ilgen, D. R., Sheppard, L. (in press). Motivation in teams. In M. Erez, H. Thierry, & U. Kleinbeck (Eds.), Motivational issues in multi-national organizations. New York: Erlbaum.

Conference Presentations

17. Ilgen, D. R. (1992, May). Team decision making research and theory. Symposium organized by the author. Annual meeting of the Society for Industrial and Organizational Psychology, Montreal.

18. Major, D. A., Sego, D. J., Hollenbeck, J. R., & Ilgen, D. R. (1992, May). Decision making in teams with distributed expertise. Presented at the annual meeting of the Society for Industrial and Organizational Psychology, Montreal.

19. Ilgen, D. R., Hollenbeck, J. R., Sego, D. J., Major, D. A., Phillips, J., & Hedlund, J. (1992, June). Team member abilities and problem solving strategy effects on team decision making outcomes and processes in teams with distributed expertise. BRG Symposium, Monterey, CA.

20. Sego, D. J., Hollenbeck, J. R., Ilgen, D. R., & Major, D. A. (1992, August). Team decision making accuracy within three different base rate conditions. Paper presented at the annual meetings of the American Psychological Association, Washington, DC.

21. Hollenbeck, J. R., Sego, D. J., Ilgen, D. R., & Major, D. A. (1992, September). Team decision making under stressful conditions: Construct validation of potential manipulations and measures from the $TIDE^2$ simulation. Paper to be presented at team conference at the University of South Florida.

22. Ilgen, D. R. (1993, June). Team decision making: Theory, method, and results. Presented at the annual Nags Head Conference on Groups and Organizational Effectiveness, Boca Raton, FL.

23. Ilgen, D. R. (1993, October). Team-level constructs of team decision making: Informity, staff validity, and hierarchical sensitivity. Presented at the annual meetings of the Society for Organizational Behavior, Virginia Beach, VA.

24. Ilgen, D. R., Major, D. A., Hollenbeck, J. R., & Sego, D. J. (1994, April). Decision making in teams: Raising an individual model to the team level. Presented as part of a symposium by E. Salas & R. Guzzo entitled, "Team decision-making in organizations: New frontiers" at the annual meeting of the Society for Industrial and Organizational Psychology, Nashville, TN.

25. Hollenbeck, J. R. (1994, July). Decision making in hierarchical teams with distributed expertise: A theory and data. Presented at the 23rd annual International Congress of the International Association of Applied Psychology, Madrid.

26. Ilgen, D. R. (1994, July). Work teams: Some critical issues. Presented at the 23rd annual International Conference of the International Association of Applied Psychology, Madrid.

27. Ilgen, D. R. & Hollenbeck, J. R. (1994, July). Decision making accuracy in four person teams with distributed expertise: Testing the predictive utility of three team-level constructs. Presented at the 1994 Symposium on Command and Control Research and Decision Aids, Naval Postgraduate School, Monterey, CA.

28. Hollenbeck, J. R., Ilgen, D. R., Tuttle, D., & Sego, D. J. (1995). Team performance on monitoring tasks: An examination of decision errors in contexts requiring sustained attention. Paper presented at the 1995 annual Meeting of the Academy of Management, Vancouver, Canada.

29. Hollenbeck, J. R., Ilgen, D. R., LePine, J. A., & Hedlund, J. (1995). The multilevel theory of team decision making: Extensions and interventions. Paper presented at the First Annual International Conference on Command and Control Research. Washington, DC: National Defense University.

30. Ilgen, D. R., & Hollenbeck, J. R. (1995). Decision making in hierarchical teams. Symposium presentation at the First Annual Conference of the Australian Organizational Psychology Association, Sydney.

31. Ilgen, D. R. (1995). Teams in work organizations: Facts and infatuations. The Keynote Address for the First Annual Conference of the Australian Psychology Association, Sydney.

32. Ilgen, D. R. (1996, April). Computational modeling: Its potential in industrial and organizational psychology. Presented as discussant in symposium at the annual meetings of the Society for Industrial and Organizational Psychology, San Diego, CA.

33. Phillips, J. M., Hollenbeck, J. R., & Ilgen, D. R. (1996, April). The prevalence and prediction of positive discrepancy creation: An application of episodic and non-episodic theories of motivation. Presented at the annual meetings of the Society of Industrial and organizational psychology, San Diego, CA.

34. Hollenbeck, J. R., Ilgen, D. R., & Weissbein, D. (1996, June). Information management and decision making following unusual events: Analysis at the team and individual level. Paper presented at the 1996 Command and Control Research and Technology Symposium, Monterey, CA.

35. Hollenbeck, J. R., Ilgen, D. R., Hedlund, J., Colquitt, J. A., & LePine, J. A., (1996, April). The multilevel theory of team decision making: Replication and extension. Presented at the annual meetings of the Society of Industrial and Organizational Psychology, San Diego, CA.

36. Hollenbeck, J. R., Ilgen, D. R., & Weissbein. (1996, April). Improving decision making in contexts requiring sustained attention. Presented at the annual meeting of the Society for Industrial and Organizational Psychology, San Diego, CA.

37. Hollenbeck, J. R., Weissbein, D., & Ilgen, D. R. (1996, August). Reducing team decision performance decrements in hierarchical teams under conditions demanding sustained attention. Presented as part of a symposium entitled, Disasters and Decisions: Fearing, Surviving, and Explaining Unusual and Terrible Events," at the annual meetings of the Academy of Management, Cincinnati, OH.

38. Ilgen, D. R., & Hollenbeck, J. R. (1996, August). Multiple contributions of computer simulations to the study of team decision making. Presented as part of a symposium entitled, "Using Computer Simulations to Study Complex Organizational Behaviors," at the annual meetings of the Academy of Management, Cincinnati, OH.

39. Ilgen, D. R. (1996, August). Cross-fertilization: A two way street. Symposium speaker at the 24th International Congress of Applied Psychology, Montreal.

40. Ilgen, D. R. (1996, April). Decision making in teams: Performance and motivational issues. One of three all conference level speakers invited to present one hour symposium-wide talks at the Applied Behavioral Science Symposium, U. S. Air Force Academy.

Invited Talks

41. Ilgen, D. R. (1991, February). Team research in the 1990s. Invited presentation for Festschrift honoring the retirement of Fred E. Fiedler, Claremont-McKenna College, The Claremont Colleges, Claremont, CA.

42. Ilgen, D. R. (1991, March). Team research in the 1990s: Making progress and avoiding past failures. Perth: University of Western Australia.

43. Ilgen, D. R. (1991, March). Team research in the 1990s. Queensland, Australia: Bond University.

44. Hollenbeck, J. R., Ilgen, D. R., Sego, D. J., & Major, D. A. (1991, June). The development of a computer-based, multi-context decision making simulation for conducting research on teams and individuals. Paper presented at the annual meetings of the Personnel and Human Resource Interest Group. West Lafayette, IN: Purdue University.

45. Ilgen, D. R. (1991, November). Team decision making research. Detroit, MI: Wayne State University.

46. Hollenbeck, J. R., Ilgen, D. R., Sego, D. J., Major, D. A., Phillips, J., Hedlund, J., & Barrett, L. (1992, February). Team decision making under stressful conditions: A program of research. Ithaca, NY: Cornell University.

47. Ilgen, D. R., Hollenbeck, J. R., & Barrett, L. (1992, September). Decision making in hierarchical teams with distributed expertise: A model, method, and some data. Presented as an invited symposium, Crew Performance Function, Armstrong Laboratory, Brooks AFB, TX.

48. Ilgen, D. R. (1994, April). Recognizing limits and expanding horizons in industrial and organizational psychology. First annual invited lecturer for the newly established annual E. J. McCormick Lecture of the School of Liberal Arts and Sciences. West Lafayette, IN: Purdue University.

49. Hollenbeck, J. R., Ilgen, D. R., Tower, S., & Waldschmidt, D. (1994). Mathematical models of team decision making: Internally analyzing theories of differentiation and integration. Paper presented at the annual meeting of the Personnel/Human Resources Research Group. Iowa City: University of Iowa.

50. Hollenbeck, J. R. (1995). Sustained attention in team contexts. Invited address. Mount Pleasant: Central Michigan University.

51. Ilgen, D. R. (1995). Decision making teams with hierarchical structures and distributed expertise. Invited address. Gold Coast, Queensland, Australia: Bond University.

52. Ilgen, D. R. (1995). Decision making teams with hierarchical structures and distributed expertise. Invited address. Hong Kong: Hong Kong University of Science and Technology.

53. Ilgen, D. R. (1996, January). Team decision making: A program of research. Invited colloquium to groups and teams faculty seminar series jointly sponsored by Carnegie-Mellon University and the University of Pittsburgh.

54. Ilgen, D. R. (1996, April). Hierarchical teams with distributed expertise. Invited colloquium, College of Business, University of Florida.

55. Ilgen, D. R. (1996, April). Decision making in hierarchial teams with distributed expertise. Invited colloquium, College of Business, University of South Carolina.

56. Hollenbeck, J. R. (1996, January). Repeated measures regression in cross-cultural research: Finding where the action is in cross-level variance structures. Invited paper presented at Hong Kong University of Science and Technology.

57. Hollenbeck, J. R. (1996, May). Improving team decision making: Lessons from the Team Effectiveness Research Laboratory. Invited talk to the Michigan Association of Industrial/Organizational Psychology, Southfield, MI.

Theses and Dissertations

58. Landis, R. S. (1992, March). The effects of team composition and incentives on team performance on an interdependent task. Unpublished masters thesis. East Lansing: Michigan State University.

59. Major, D. A. (1992, August). Decision making at the individual and team levels: Moderators of the effects of cognitive frames and risk taking. Unpublished dissertation. East Lansing: Michigan State University.

60. Hedlund, J. (1993, December). Computer-mediated versus face-to-face communications in hierarchical team decision making. Unpublished masters thesis. East Lansing: Michigan State University.

61. Quinones, M. A. (1993, December). Pre-training context effect: Training assignment as feedback. Unpublished dissertation. East Lansing: Michigan State University.

62. Sego, D. J. (1994, August). Role formation within hierarchical decision making teams with distrutive expertise: A role expansion model. Unpublished dissertation. East Lansing: Michigan State University.

63. Barrett, L. (1996, December). Effect of simultaneous versus sequential display of visual information on decision making accuracy: Moderating effects of decision context. Unpublished dissertation. East Lansing: Michigan State University.

64. Phillips, J. (1996). Antecedents and consequences of leader utilization of staff information in decision making teams: Address a leadership dilemma. Unpublished dissertation. East Lansing: Michigan State University (Final oral defense scheduled for January 17, 1997).

Other

65. Development of a Task and Laboratory Area for Team Decision Making Effectiveness Research Laboratory at Michigan State University.